



This article was published in an Elsevier journal. The attached copy is furnished to the author for non-commercial research and education use, including for instruction at the author's institution, sharing with colleagues and providing to institution administration.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>

Research paper

Release from interference in auditory working memory for pitch

Dennis T. Ries^{a,*}, Jeffery J. DiGiovanni^b

^a Auditory Perception Laboratory, W241 Grover Center, School of Hearing, Speech, and Language Sciences, Ohio University, Athens, OH 45701, USA

^b Auditory Psychophysics and Signal Processing Laboratory, W151a Grover Center, School of Hearing, Speech, and Language Sciences, Ohio University, Athens, OH 45701, USA

Received 1 February 2007; received in revised form 23 April 2007; accepted 24 April 2007

Available online 8 May 2007

Abstract

The purpose of this study was to quantify the effect of interpolated tones upon a pitch standard held within auditory working memory through measurement of the difference limen (just noticeable difference) for frequency and the usefulness of “Where” cues to ameliorate the interference produced by these intervening stimuli. To this end, we measured the degree to which tones, containing identical and disparate localization cues, presented within the retention interval altered differential sensitivity for frequency via the method of constant stimuli. The difference limen for frequency nearly tripled when tones were presented within the retention interval and sound localization cues produced a significant partial release from interference within the short-term pitch store. Interference produced by “Where” cues ranged from 4.0 to 5.2 Hz. These findings indicate that there is a possible integrative use of the “What” and “Where” pathways in forming and maintaining pitch information within the pitch array within auditory working memory.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Pitch; Difference limen; Auditory; Working memory; Frequency

1. Introduction

It has long been known that the small difference in frequency that can just be detected (i.e., the difference limen) for pitch slowly increases as the temporal gap between a standard and comparison tone increases (Harris, 1952; Koester, 1945). In the late 1960s and early 1970s, a series of pitch memory experiments examined the role of interference upon pitch perception (Bull and Cuddy, 1972; Deutsch, 1970a,b, 1972a; Elliot, 1970; Massaro, 1970; Wickelgren, 1966, 1969). The independent studies conducted by Deutsch and Massaro focused on the extent to which intervening (interpolated) sounds placed within a temporal retention interval lasting several seconds (e.g.,

≈5 s) disrupts the ability of listeners to identify whether two target sounds (i.e., a standard and a comparison tone) were the same or different. The results of these investigations revealed that intervening tones produced the greatest disruption of pitch memory for tonal targets whereas interpolated speech or noise stimuli had little effect. In addition, Deutsch (1970a, 1972a,b) observed that the degree of interference was dependant upon the relationship of the intervening frequencies with those of the standard and comparison. Later studies (Deutsch, 1974a; Deutsch and Feroe, 1975; Semal and Demany, 1991, 1993) have expanded upon these earlier findings regarding the close relationship between the pitch of the interpolated tones to that of the stored standard and the magnitude of memory disruption. In addition, Pechmann and Mohr (1992) have found that extensive musical training can reduce the deleterious effects of intervening interference upon retention of pitch within working memory (see also Berti et al., 2006). Based on results such as these, researchers have concluded that the interactions were occurring within

* Corresponding author. Address: School of Hearing, Speech, and Language Sciences, W221 Grover Center, Ohio University, Athens, OH 45701, USA. Tel.: +1 740 593 1420; fax: +1 740 593 0287.

E-mail addresses: ries@ohio.edu (D.T. Ries), digiovan@ohio.edu (J.J. DiGiovanni).

a short-term store specially allocated for pitch memory (e.g., Berti et al., 2006; Deutsch, 1972a, 1999; Pechmann and Mohr, 1992; Semal and Demany, 1991, 1993).

Results from further investigations elucidate that basic aspects of auditory stimuli (e.g., pitch) are processed to some degree via separate cortical networks (Alain et al., 2001; Arnott et al., 2005; Clément et al., 1999; Deutsch, 1999; Maeder et al., 2001; Vuontela et al., 2003). These networks include those involved in the function of working memory (Anourova et al., 1999; Arnott et al., 2005; Martinkauppi et al., 2000; Vuontela et al., 2003). Those processes involved in short-term recall and comparison of auditory percepts within working memory are linked primarily to the frontal (Baddeley, 1986, 1992) and prefrontal (Braver et al., 1997; Carlson et al., 1998; Casey et al., 1995; Martinkauppi et al., 2000) areas of the cortex that engage systems linked with central executive function and control of attention. Conversely, the retention of information within working memory primarily takes place in the superior temporal gyrus and parietal cortex (Arnott et al., 2005; Cowan et al., 2005). However, the amount of retained information that is available at any one time is dependant upon attention-based processes controlled in the frontal and prefrontal areas within the cortex (Cowan et al., 2005) and this plays a role in one's ability to keep a highly accurate representation of a percept alive within working memory in the presence of interference (Engle et al., 1999). These storage and processing components of the working memory architecture are further subdivided such that basic attributes of auditory perception are segregated broadly into categories of "What" and "Where" (e.g., Alain et al., 2001; Arnott et al., 2005; Deutsch and Roll, 1976; Rauschecker and Tian, 2000) with the "What" category being subdivided further into memory modules for specific auditory attributes such as duration and loudness (Deutsch, 1999). This differentiation of auditory coding mechanisms, in conjunction with the attention-based processes, may permit a listener to retain more accurate information about the pitch of a target sound within auditory working memory (AWM) when the "Where" cues of later occurring interferers differ from that of the standard stored within short-term memory.

Release (disinhibition) of interference within AWM for pitch need not solely rely upon "Where" cues (e.g., Berti et al., 2006; Deutsch, 1970b, 1972b, 1975; Deutsch and Feroe, 1975; Pechmann and Mohr, 1992), but, nonetheless, localization information has proved sufficient to produce the phenomena (Deutsch, 1978a; Kallman et al., 1987). That is, improvements in the ability to retain a reasonably accurate rendition of a target pitch in the memory store has been shown to be related to the interferer-to-standard pitch relationship (e.g., Deutsch and Feroe, 1975), musical training (e.g., Berti et al., 2006; Pechmann and Mohr, 1992), as well as to dichotic presentation of the target and interfering tones (Deutsch, 1978a; Kallman et al., 1987). Deutsch (1978a) found that subjects produced fewer errors for the identification of a semitone difference between the

standard and comparison stimuli when the test tones and intervening interference tones were presented to opposite ears. The decrease in detection errors in the dichotic test tone/interfering tone conditions was small (roughly 3–5%) but statistically significant.

However, this improvement is tempered somewhat when the ear (ipsilateral or contralateral with respect to the ear receiving the test tones) presented with the interpolated tones changes randomly from trial to trial (Kallman et al., 1987). Furthermore, there is evidence that binaural differences in stimulus attributes does not always lead to an increase in pitch perception performance (e.g., Deutsch, 1974b, 1988).

Deutsch performed a series of experiments investigating what she termed the *octave illusion* in which sequential dichotic presentations of tonal sequences produced an inaccurate perception (Deutsch, 1974b, 1978b, 1980, 1988, 2004; Deutsch and Roll, 1976). In the original report of the illusion (Deutsch, 1974b), she found that when high and low tones separated by an octave were alternately presented to separate ears (e.g., high tone to the right ear and low tone to the left ear for 0.25 s, followed by the opposite presentation for the next 0.25 s, with the sequence repeating over a span of 1.5 s) a majority of listeners reported hearing "a single tone oscillating from ear to ear, whose pitch also oscillated from one octave to the other in synchrony with the localization shift" (p. 307) and none of listeners perceived the correct sequence of events. The lateralization shift followed the location of the higher frequency tone (Deutsch, 1980; Deutsch and Roll, 1976) provided that the intensity of the lower frequency stimulus did not surpass that of the higher frequency signal by 12 dB or more (Deutsch, 1980, 1988). The perceived pitch of the tone was determined by the frequency presented to the dominate ear at a given point in time (Deutsch, 1980; Deutsch and Roll, 1976). The illusion holds up well for alternation rates between approximately 0.05 and 0.5 s with subjects obtaining correct percepts (no illusion) over 80% of the time when the switching rate was 2.0 s (Zwicker, 1984), as cited in Deutsch (2004). Deutsch (1975), as cited in Deutsch (2004) and Deutsch and Roll (1976) proposed a model in which the illusion arises from separate "What" and "Where" cue processes. Thus, from the results of the *octave illusion* and dichotic interferer test tone studies (e.g., Deutsch, 1978a), it appears that dichotic cues can be beneficial or deleterious to accurate perception dependant upon the characteristics of the cues.

To our knowledge only one study on memory for pitch has addressed the effect that interference has upon the resolving power of the auditory system for pitch. That is, most of the earlier research in this area (e.g., Deutsch, 1970a,b, 1972a, 1978a,b; Massaro, 1970) demonstrated that the ability of listeners to correctly identify a fixed frequency difference (e.g., a semitone difference) was diminished by the presence of interferers, but this work did not address the effect of such intervening sounds upon a person's absolute sensitivity for changes in frequency. Elliot (1970), however,

found that the presence of a single interpolated tone placed in close temporal proximity to either a standard tone or a comparison tone roughly doubled the difference limen (DL) with the effect being slightly more pronounced when the intervening tone was placed close to the standard. That is, the smallest frequency difference that could just be detected needed to be approximately twice as large on average (e.g., 6 Hz vs. 3 Hz) when the interpolated tone was present and needed to be slightly larger (e.g., 7 Hz vs. 5 Hz) when the intervening tone was placed closer to the standard to be retained. The findings reported by Elliot (1970) show that the degree of disruption produced by interpolated tones upon a pitch standard stored within AWM can be quantified through a frequency discrimination task.

The frequency discrimination results reported by Elliot (1970) for the conditions containing intervening interferers, however, are limited by two main factors. First, her measures only made use of one interpolated tone within the retention interval whereas most other pitch memory experiments using interferers have employed 4–8 intervening stimuli (e.g., Berti et al., 2006; Deutsch, 1970a,b, 1972a, 1978a,b; Kallman et al., 1987; Massaro, 1970; Pechmann and Mohr, 1992; Semal and Demany, 1993). The amount of interference produced by interpolated tones has been shown to increase as the number of stimuli within the retention span increases (Deutsch, 1970, as cited in Deutsch, 1999; Wickelgren, 1966, 1969). Second, the temporal separation between the standard and interpolated tone or the interpolated tone and the comparison tone of 50 ms in Elliot's study was not sufficient to control for the effects of forward masking. Furthermore, while it is known that "Where" cues can lessen the effects of interference for pitch memory (e.g., Deutsch, 1978a; Kallman et al., 1987) this release from interference has not been quantified in terms of the resolving power (i.e., measurement of the DL) of the auditory system.

In this light, the objective of our study was to differentiate the effects of acoustic interferers upon the capacity to resolve frequency differences by judging whether a later occurring comparison tone was higher or lower in frequency than the standard maintained within AWM. The hypotheses of this study are as follows: (1) The introduction of four interpolated tones within the retention interval during a delayed comparison task will disrupt the accuracy of a standard stored within AWM producing an increase in the DL over that measured when no intervening stimuli are present; (2) the introduction of "where" cues via interaural intensity or phase differences within the interpolated tones will produce a release from interference resulting in a lower DL than that measured in the presence of interferers that are devoid of these spatial indicators.

2. Materials and methods

2.1. Subjects

Eight subjects from the Ohio University student, faculty, and staff population participated in the study. All subjects

had prior experience in psychoacoustic experiments. They were recruited by personal request and received no compensation for their participation. All subjects had pure-tone, air-conducted thresholds of less than 15 dB HL at the octave frequencies from 250 to 8000 Hz. The subjects were not selected based upon any previous musical training, although subjects S2 and S4 had 6–7 years of musical instruction prior to their entering college and subjects S1, S5, and S6 had limited (<1 year) musical training. Subjects S3, S7, and S8 had no prior musical training. Only subjects S2 and S4 reported that they still played a musical instrument and both reported that they play their instruments on a sporadic basis. All subjects provided informed consent in accordance with the approval of the study by the Institutional Review Board at Ohio University.

2.2. Stimulus characteristics and presentation

All stimuli were generated and output via System 3 hardware (Tucker Davis Technologies, Alachua, FL) controlled by a Pentium 4 computer (Dell, Round Rock, TX) running Matlab (The Mathworks, Natick, MA) and RPVD (Tucker Davis Technologies) software. The signals were verified electrically via a dynamic signal analyzer (Stanford Research Systems, Sunnyvale, CA) as well as a digital storage oscilloscope (Tektronix, Richardson, TX) and acoustically using a sound level meter and 2cc acoustic coupler (Brüel and Kjær, Naerum, Denmark). The signals were presented binaurally through ER-2 insert earphones (Etymotic Research, Elk Grove Village, IL) to subjects seated with a double walled, sound attenuating booth (Industrial Acoustics Corporation, Bronx, NY).

Stimuli were sinusoids with durations of 200 ms, including 20 ms cosine-squared onset and offset ramps generated at a sampling rate of 24,414 Hz and low pass filtered at 12,000 Hz. The offset of the standard tone (first sinusoid presented on a given trial) was separated from the onset of the comparison tone (last sinusoid presented on a given trial) by 4800 ms. The standard tone frequency was roved randomly over a range of 80 Hz centered at 435 Hz (range = 395–475 Hz) and presented at a level of 60 dB SPL. The frequency of the comparison tone was determined randomly for each trial from a set of 11 frequency differences set relative to that of the standard. The frequency of the comparison was selected from a 100 Hz range (0 Hz, ± 10 Hz, ± 20 Hz, ± 30 Hz, ± 40 Hz, and ± 50 Hz) for runs containing intervening stimuli and from a 40 Hz range (0 Hz, ± 4 Hz, ± 8 Hz, ± 12 Hz, ± 16 Hz, and ± 20 Hz) for the no interference condition (hereafter referred to as NoINT). The frequency range of the comparison for the no interference condition was different from that of the conditions containing interference so as to obtain several points of performance throughout the psychometric function that were not limited by floor and ceiling effects. For example, if the range of the no interference condition were identical to that used in the interference conditions, subjects likely would produce only one data

point, in this case at ± 10 Hz, between chance performance and near perfect performance and thereby limiting the accuracy of the sigmoid curves fit to the data. Therefore, the spacing of the frequency differences, determined from pilot data, was different between the conditions containing interpolated tones and the NoINT condition. This is not a concern for later analysis as the measure of interest is the frequency difference that is detectable 75% of the time for each condition and not a comparison of performance at identical frequency separations between the standard and the comparison. The runs containing intervening stimuli included four additional tones while the control runs were comprised of only the standard and comparison stimuli.

When intervening stimuli were present, the onset of the first began 1000 ms after the offset of the standard tone with 300 ms interstimulus intervals placed between the four intervening tones and a 2100 ms interval placed between the offset of the last intervening tone and the onset of the comparison tone (see Fig. 1). The intervening stimuli consisted of four separate sinusoids covering a frequency range of 183–691 Hz that correspond to notes of an equal tempered scale using international pitch whereby 435 Hz is the note A (Deutsch and Feroe, 1975). The frequency of each intervening stimulus was never to equal that of either the standard or comparison tone. This set of parameters defined the interpolated tones (hereafter referred to the INT condition) used to measure the effect of introducing intervening stimuli within the retention interval upon the DL for the target tones. All of the interfering tones in the study were equated for loudness except as noted in the level difference conditions below.

Two separate methods were used to produce four additional conditions in which the interferers differed in intensity or starting phase from that of the standard/comparison pair. In each of these additional interference conditions, the characteristics of the target tones were unaltered from the previous two conditions and were perceived at the mid-line of the head. Two of these four additional conditions were produced by adjusting the amplitude of the interpolated tones. In one of these conditions (hereafter referred to as ILD for interaural level difference), the level of the interpolated tones presented to the left ear were

reduced by 12 dB relative to the right ear producing a perception that the intervening stimuli were present only in the right ear. The amplitude of the right ear stimulus was identical to those used in the INT condition. In the other level difference condition (hereafter referred to as D12dB), the level of the interpolated tones was reduced by 12 dB in both ears relative to the level of these the same stimuli in the INT condition. In addition, two conditions in which the starting phase of the interferers differed from that of the target tones were presented. In one of these conditions (hereafter referred to as IPD180), an interaural phase difference was introduced in which the starting phase of the interpolated tones presented to the right ear was 180° out of phase with that presented to the left ear (the left ear starting phase was 0°) producing an more diffuse percept for the location of intervening tones compared to that of the unaltered target stimuli. The other phase difference condition (hereafter referred to as IPD90) introduced starting phase differences that produced a consistent 90° lateralization of the interferers to the right calculated based on the equations provided by Kuhn (1977). The ILD, IPD180, and IPD90 cues were employed to evaluate whether dichotic (i.e., “Where”) cues introduced to the interpolated tones presented to each ear produce a release for interference as has been reported in earlier investigation (Deutsch, 1978a,b; Kallman et al., 1987) which simply presented the target tones and the interpolated tones to opposite ears. The D12dB condition was presented to determine whether any improvements that may be obtained in the ILD condition were due to a “Where” cue or a simple intensity difference between the target tones and the interpolated tones.

2.3. Procedure

Measures of subject performance were obtained using the method of constant stimuli combined with a variation of the timing skeleton used by Deutsch (e.g., 1970a, 1970b, 1972a). In this procedure, the standard sound remains constant while the comparison signal is variable in value. The value of the comparison for any given trial was selected randomly from a predetermined range (see stimulus characteristics and presentation above). The subject was required to select whether the comparison was higher or lower in pitch than the standard by touching the area on a touch screen monitor marked “H” for higher or “L” for lower. Higher and lower frequency signals were employed to control for the response bias that likely would occur if the response choices were same or different. That is, in our implementation of the procedure, if the response format was simply same or different the subject would be right 100% of the time on the difference intervals simply by selecting different on every trial. The higher and lower response format lowers the limits of the upper performance level to 50% correct should the subject select the identical response (i.e., always selects higher or always selects lower) on every trial.

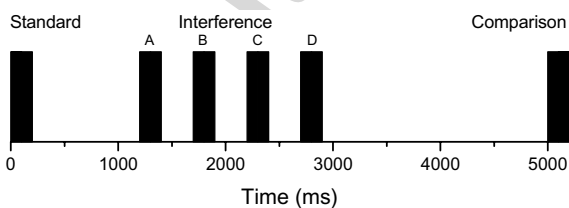


Fig. 1. Timing skeleton employed in the study depicting the temporal configuration of the standard, comparison, and the four (A, B, C, and D) intervening interference stimuli. The interfering stimuli did not appear in the NoINT condition, but were presented in all other conditions. The level (amplitude) or starting phase of the interferers differed in some conditions as described in the text.

Each run in an experiment included 10 presentations of each frequency difference with 20 presentations being collected for the no difference condition. This allowed for an equal number of trials to be collected for a given absolute frequency difference (e.g., 20 responses each for 0, 10, 20, 30, 40, and 50 Hz separations in frequency). The zero frequency difference condition was used to estimate response bias for each listener. As described above, eleven different comparison values were employed (5 above the value of the comparison, 5 below the value of the comparison, and one of equal value). Thus, each run involves 120 trials (10 presentations \times 10 different comparisons plus 20 presentations for the 0 Hz frequency difference) and each participant completed three runs for both the control and interference conditions, resulting in 60 judgments per participant for each absolute frequency difference. Each run took approximately 35–45 min to complete, including breaks, which were offered after every 20 responses. Each participant received practice runs using the control condition until such time as they felt comfortable with the task and the response format prior to beginning formal collection of the data. The order of the runs was randomized across conditions.

3. Results

3.1. Intervening stimuli impair a listener's ability to discriminate differences in pitch

Sigmoid functions were fit to the data of each subject anchoring the lower end of the function to chance performance (i.e., 50%). That is, 50% correct performance was inserted as the value in the zero frequency difference condition as the scores determined for this point represented response bias and not true performance since the percent correct values in this condition were calculated by arbitrarily assigning a “higher” response as correct and a lower response as incorrect. Table 1 lists the response bias per condition for each listener (bias results presented below). The difference limen (DL) for pitch (i.e., fre-

quency) was taken as the point on the resulting function that corresponded to 75% correct identification. The DLs across subjects were analyzed by way of a repeated measures analysis of variance (ANOVA) with the six conditions serving as factors (independent variables). The Geisser-Greenhouse Adjustment was applied to correct for violation of sphericity. Overall, there is a significant difference for condition ($F(5,35) = 9.80, p < 0.05$) and as such post hoc comparisons between the NoINT condition and those containing interpolated tones were evaluated using Fisher's LSD multiple comparison test (Ryan, 1959). As can be seen from the results depicted in Fig. 2, the ability to discriminate a difference in pitch between the standard and comparison stimuli was diminished (DL increased) significantly (Fischer LSD, $p < 0.05$) in all conditions containing interfering tones (i.e., conditions INT, ILD, D12dB, IPD180, and IPD90). That is, a greater change in frequency (y -axis) was required for the subjects to discriminate a difference in frequency correctly 75% of the time for any of the conditions (x -axis) containing interference. The average increase in the DL for pitch was 8.6 Hz across all conditions containing interfering tones with the average DL in the INT condition (no level or phase cues) being almost triple (2.98 times) that observed in the NoINT condition. It should be noted that some listeners (e.g., S1) were more affected by the intervening tones across most conditions than were others (e.g., S4). These results indicate that sounds occurring within the retention interval interfere with the ability of listeners to maintain an accurate auditory object within AWM for the standard stimulus.

3.2. Interaural level and phase differences produce a partial release from memory interference

Further examination of the results in Fig. 2 reveal that the DL for pitch decreases (i.e., improves) in the presence of intervening sounds when the intervening sounds can be perceptually segregated from the standard and comparison stimuli by means of ILDs (ILD condition) or IPDs

Table 1
Percentage of higher responses when the target tones were of identical frequency from each subject for each condition as an indicator of response bias

Subject	Condition						Subject average
	NoINT	INT	ILD	D12dB	IPD180	IPD90	
S1	50.0	45.0	65.0	35.0	46.7	58.3	50.0
S2	45.0	50.0	60.0	56.7	45.0	43.3	50.0
S3	50.0	48.3	35.0	43.3	51.7	48.3	46.1
S4	45.0	55.0	62.5	25.0	65.0	60.0	52.1
S5	86.7	48.3	28.3	38.3	40.0	48.3	48.3
S6	50.0	46.7	36.7	35.0	53.3	30.0	42.0
S7	78.3	65.0	78.3	61.7	53.3	83.3	70.0
S8	70.0	71.7	60.0	75.0	71.7	83.3	72.0
Condition average	59.4	53.8	53.2	46.3	53.3	56.9	53.8

Average bias for responding higher for each condition across subjects is provided in the bottom row and average bias for each subject across condition is provided in the rightmost column. The overall average bias across all conditions and all subjects is provided by the entry in the lower, right-hand corner. Chance performance (i.e., no bias) for the task was 50%.

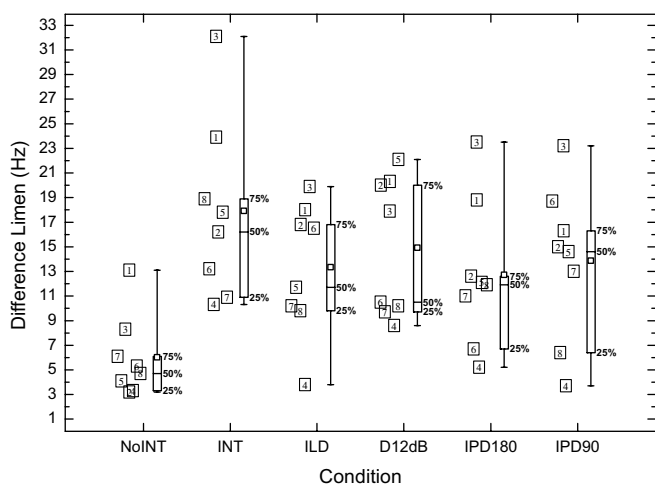


Fig. 2. Difference limen for frequency (y -axis) for each interference condition (x -axis). The box plot for each condition illustrates the median (line labeled as 50%), the 25th and 75th percentiles for the data (lower and top edges of the box, respectively), and the mean of the data across subjects for each condition (open square). The end points of the whiskers extending from each box demarcate the 1st and 99th percentiles of the data for each condition. Individual data points per condition are represented by the square symbols just to the left of the corresponding box and whiskers plot. The number within each square indicates which participant produced a given data point within each condition (e.g., S1's results for each condition are represented by the boxes containing the number 1). The condition abbreviations are as follows: NoINT, no intervening stimuli within the retention interval; INT, intervening stimuli contain no diotic or dichotic phase or level cues to set them apart from the target tones; ILD, the left ear intensity of the intervening stimuli in the left ear is 12 dB less than that of the right ear and the right ear intervening stimuli are equal in loudness to the target tones; D12dB, the intensity of the intervening stimuli in both ears is approximately 12 dB less than the target tones; IPD180, the starting phase of the intervening stimuli was 0° in the right ear and 180° in the left ear; IPD90, the starting phases of the intervening stimuli produced a perceived 90° lateralization to the right.

(IPD180 and IPD90 conditions). The size of the effect varied from subject to subject, but the average improvement in DL across subjects was similar (4.0–5.2 Hz) for these conditions (i.e., ILD, IPD180, and IPD90) relative to the INT condition. A smaller average improvement in the DL of 3.0 Hz was obtained when comparing the D12dB and INT conditions.

A closer examination of the individual data reveals that all subjects exhibited an increase in their DL from the NoINT to the INT condition. Six of 8 subjects in the ILD, IPD 180, and IPD 90 conditions and 5 of 8 in the D12dB condition exhibited lower DLs when compared to the INT condition. In addition, when comparing performance in the INT condition to the conditions containing dichotic phase or level information for the interpolated tones (i.e., the ILD, IPD180, and IPD90 conditions), S6 performed opposite of the average trend except in the IPD 180 condition, S7 exhibited little change in his/her DL across all conditions containing intervening tones, and S2 had no improvement in his/her DL in the ILD condition, but exhibits lower DLs in the two dichotic phase conditions (i.e., IPD180 and IPD90). Finally, subjects S4

and S7 had only minor (<1.8 Hz) decrease in their DL for the D12dB conditions whereas subjects S2 and S5 had increases in their DL in this same condition of 3.8 and 4.3 Hz, respectively.

As the overall F -statistic with Geisser-Greenhouse Adjustment was significant (see Section 3.1), post hoc pairwise comparisons were evaluated using Fisher's LSD multiple comparison test (Ryan, 1959). The results of the post hoc analyses indicate that (1) the DLs obtained across all subjects for the ILD and IPD interference conditions (IPD180 and IPD90) were significantly smaller than those measured in the INT condition (Fisher LSD, $p < 0.05$); (2) the DLs measured across all subjects in the D12dB interference condition were significantly larger than those obtained for the ILD, IPD90, and IPD conditions (Fisher LSD, $p < 0.05$), but were not significantly different from the INT condition (Fisher LSD, $p > 0.05$). These results indicate that location cues produced by interaural (dichotic) level or phase differences between ears reduce the effect of intervening interference within AWM significantly, but that a simple diotic difference (i.e., interpolated tone amplitude 12 dB below the standard/comparison pair level, binaurally) between the standard/comparison pair and the intervening stimuli is not sufficient to facilitate this effect.

Table 1 shows the percentage of the time that a subject responded higher when the target tones had identical frequencies per subject and condition. As can be seen in this table, there was little overall bias across conditions with only 53.8% (see rightmost cell of the bottom row) of the overall response being higher when the target tones were of identical frequency. Across subjects the average percent of higher responses was near 50.0% (chance performance on this task) for each condition (see condition average row on the table) and ranged from 46.3% in the D12dB condition to 59.4% in the NoINT condition.

Furthermore, 6 of the 8 subjects (i.e., S1–S6) responded within 8.0% of chance performance across conditions (see subject average column on the table). Subjects (S7 and S8), however, were clearly biased towards responding higher overall. In addition, Subject S5 had the greatest change in bias across conditions ranging from 28.3% in the ILD condition to 86.7% in the NoINT condition. The D12dB condition was the only instance in which the majority (5 out of 8) subjects were biased towards responding lower. In this case, 4 of the 5 subjects responding lower in this condition exhibited a bias of greater than 11% (range 11.7–25%) toward responding lower. In contrast, 5 of 8 subjects in the ILD and IPD180 conditions were biased towards responding higher. In all three cases (majority low response bias in D12dB condition, majority higher response bias in the ILD and IPD180 conditions), the effects was essentially negated by the minority responses biased in the opposite direction. It is interesting to note that the only condition in which large individual response biases in one direction (high or low) are not counterbalanced by other individual biases in the opposite direction is in the NoINT condition.

The response bias results were analyzed by two separate repeated-measures ANOVAs, one with condition as the factor and the other with subject as the factor. The later analysis was conducted due to the clearly different bias exhibited by subjects S7 and S8. The results of the condition bias analyses indicate that response bias did not differ significantly across conditions ($F(5, 35) = 1.15$, $p = 0.35$). The results of the subject bias analysis with a Geisser-Greenhouse correction, however, indicate that there was a significant difference in response bias across subjects ($F(7, 35) = 5.35$, $p < 0.05$). As the overall F -statistic was significant, post hoc analyses were completed using Fisher's LSD multiple comparison test. The pairwise comparisons indicate that subjects S7 and S8 tended to respond higher significantly (Fischer LSD, $p < 0.05$) more often than the other six subjects. Overall, the results presented above indicate that differences in the DLs measured are a result of differences in the ability of subjects to resist the deleterious effect of the interpolated tones and that the method used controlled well the effects of response bias upon performance.

4. Discussion

The results obtained when comparing the no interference conditions to those containing interpolated tones support the notion that the intervening stimuli placed within the retention interval produce a significant degradation of a pitch standard held within memory. The extent of this degradation can be estimated through measurement of DLs for frequency. The DLs from the no interference condition in the present experiment are comparable to those reported by Harris (1952) in his study on the decline of frequency discrimination over time and roughly double (approximately 6 Hz vs. 3 Hz) that reported by Elliot (1970). In addition, the DL in the interference condition without "Where" cues (INT condition) is approximately three times as large (approximately 18 Hz vs. 6 Hz) as that reported by Elliot (1970). The difference in the DLs measured in the presence of interpolated tones between our study and Elliot's is most likely due to the differing number of intervening tones employed in the two investigations as the degree of disruption of the standard within working memory increases concurrent with increases in the number of tonal stimuli placed within the retention interval (Deutsch, 1970, as cited in Deutsch, 1999). The present study employed four interferers while Elliot's used one and therefore we would expect the DL in the current investigation to be greater. Therefore, Elliot's finding that the DL doubles in the presence of a single interpolated tone and the present finding that this same measure triples when presented four intervening stimuli is in accordance with the finding of Deutsch (1970, as cited in Deutsch, 1999).

In contrast, the reason for the disparity in the DLs in the no interference conditions between the two studies is less clear. The most plausible explanation would be differences in the methods employed. While both investigations required subjects to respond "high" or "low" after each

trial, Elliot produced her functions based upon the percent of "higher" responses while the present study used the percent of correct responses at each absolute frequency difference to produce the functions. As a result, the functions in the present study can span from chance performance at 50% correct to perfect performance (100% correct) whereas Elliot's functions could span from 0% to 100% correct producing a difference in the curves fit to the data in each study leading to different estimates of the DL. This would likely have an effect upon the curves fit to the data for the interpolated tone conditions as well, providing an additional reason for the differences encountered in the interpolated tone condition between the two studies.

Furthermore, the procedure used in the present study was designed to help control for the effects of response bias upon the results. For a given condition in the present study, some listeners would tend to respond "higher" if unsure whereas others would tend to respond "lower" if unsure as evidenced by their performance in the trials containing no frequency difference (see Table 1). Elliot's results clearly were affected by response bias as evidenced by the disparate number of points above and the 50% point on the representative results depicted in her study (1970), however, it is unclear the extent to which this effected her measurement of the DL for pitch. The overall response bias in the present study was small and did not differ appreciably across condition. Interestingly, Massaro (1970) reported a significant difference in bias across conditions, but the differences between the methods and measures employed make a comparison between his investigation and the present results tenuous at best. Regardless, the outcomes of the present study follow the general trend of those reported earlier by Elliot (1970).

Moreover, the present findings show that perceived spatial separation between the target tones and the interpolated sinusoids provides a partial release from interference within AWM. Disinhibition of interference within pitch memory through presentation of the interpolated tones in one ear and delivery of target tones to the other have been demonstrated previously (Deutsch, 1978a; Kallman et al., 1987). The present study, however, produced "Where" cues by providing dichotic phase or level cues while still presenting the interpolated tone to both ears in order to provide a closer simulation of real-world localization information. The results show that there is a small, yet significant improvement in the DL of approximately 4–5 Hz when interaural level or phase cues are present. The finding that the D12dB condition failed to produce to significant reduction in the DL is in accordance with those of Semal and Demany (1993) who reported that intensity differences amongst the interpolated tones has hardly any impact upon the extent of degradation of a pitch standard held within working memory. Overall, the outcomes across conditions confirm those reported by Deutsch (1978a) as well as Kallman and Colleagues (1987), expanding upon these earlier findings to encompass "Where" cues produced through binaural

presentation of phase and level differences between ears. In addition, while these earlier studies of pitch memory involving spatial separation of the target tones and interpolated stimuli measured the effects of interference by means of percent correct (Kallman et al., converted these measures to d' values), the present outcomes quantify the extent of degradation of the pitch standard through measurement of the DL. In any case, it is clear that perceived differences in the location of the intervening tones relative to the targets lessens the interaction of the interpolated tones with the information about the standard that is actively retained within the pitch store (Deutsch, 1978a).

These results support a hypothesis wherein the auditory system uses the binaural differences provided in the intervening stimuli to disassociate the interpolated tones from the standard and comparison sounds through perceived spatial separation. Active retention of a specific percept within a designated short-term memory store (e.g., pitch storage module) is thought to be under the purview of the attention-based processes of working memory (e.g., Cowan et al., 2005; Demany et al., 2004; Deutsch, 1978a, 1999; Engle et al., 1999). It has been theorized that percepts for specific auditory traits (e.g., pitch, loudness, duration) are stored in semiautonomous parallel arrays whose information is recombined upon retrieval (Deutsch, 1999) and using spatial separation across sequential tones to disassociate one tonal percept from another (Deutsch, 1978a,b; Kallman et al., 1987). These data fit Deutsch's (1972a) three-dimensional theory of pitch memory, which includes time, pitch, and strength of response, if a fourth dimension of spatial location were included. The information regarding sound source location would be processed simultaneously and in effect further strengthen those percepts being of similar pitch that originated from like locations by summation of pitch strength across time and space. Furthermore, the linking of pitch and location information would be greatest when the location of the target tones and interpolated stimuli are predictable, as suggested by the results of Kallman and Colleagues (1987), presumably by reducing the demand for attention-based resources. That is, the "Where" information acts to set place markers within the scene held in AWM (Darwin and Hukin, 1999) thereby freeing limited attention-based resources for the task of pitch preservation. Furthermore, the ability of AWM to link across arrays before retrieval is limited (Deutsch, 1999) and therefore the contribution to the strength of the pitch store is tempered, especially when interpolated information of similar type (e.g., tones) comes from random spatial locations (see Kallman et al., 1987).

Spatial location, however, is not the only way in which the pitch store can be strengthened. Pechmann and Mohr (1992) as well as Berti and Colleagues (2006) found that trained musicians are more resistant to the interference from interpolated tones than are non-musicians or those with more limited musical training. These results indicate that the pitch store in these listeners is shielded from interference through enhanced segregation of the pitch store

from the other arrays within AWM. Two of the subjects in the present experiment (S2 and S4) had musical training that exceeded 6 years whereas the remaining subjects all had less than one year on musical training. Comparing the individual results of subjects S2 (7 years musical training) and S4 (6 years musical training), it is readily evident that these two subjects have the lowest DLs when no interference is present, but that this trend only holds for S4 when interference is present. Subject S2 had higher than average DLs on the ILD and D12dB conditions relative to the INT condition and near average performance on the IPD90 and IPD180 conditions. In addition, the subjects with limited (<1 year) musical training do not perform remarkably better than those without any musical training. It remains unclear from the present data that any prior musical training influenced the results in the conditions containing interpolated tones in this study.

Overall, the present study provides further evidence that AWM for pitch for two target sounds separated in time is affected negatively when interpolated tones are placed with the retention interval. This was evidenced by an almost threefold increase in the DL for frequency when interferers with the same perceived spatial location as the target tones were present. In addition, spatial separation of the interpolated tones from the target stimuli via introduction of differences in interaural level or phase lead to a small, but significant, decrease in the DL for frequency when compared to a condition containing intervening sounds that were perceived as coming from the same location. It appears that AWM can make limited use of predictable "Where" cues (e.g., sounds emanate from consistent locations across trials within a run) to aid in shielding a standard held actively within the pitch store from some of the otherwise deleterious interference produced by interpolated tones. Current investigations are underway to assess further aspects of pitch memory using the DL as a way to quantify the extent of deterioration of the standard within the pitch memory array as well as to study similar effects in other arrays (e.g., loudness, duration) within AWM.

References

- Alain, C., Arnott, S.R., Hevenor, S., Graham, S., Grady, C.L., 2001. 'What' and 'where' in the human auditory system. *Proc. Natl. Acad. Sci. USA* 98, 12301–12306.
- Anourova, I., Rämä, P., Alho, K., Kivusalo, S., Kahnari, J., Carlson, S., 1999. Selective interference reveals dissociation between memory for location and pitch. *Neuroreport* 10, 3543–3547.
- Arnott, S.R., Grady, C.L., Hevenor, S.J., Graham, S., Alain, C., 2005. The functional organization of auditory working memory as revealed by fMRI. *J. Cogn. Neurosci.* 17, 819–831.
- Baddeley, A.D., 1986. *Working Memory*. Clarendon, Oxford, England.
- Baddeley, A.D., 1992. Is working memory working? *Quart. J. Exp. Psychol.* 44, 5–28.
- Berti, S., Münzer, S., Schröger, E., Pechmann, T., 2006. Different interference effects in musicians and a control group. *Exp. Psychol.* 53, 111–116.
- Braver, T.S., Cohen, J.D., Nystrom, L.E., Jonides, J., Smith, E.E., Noll, D.C., 1997. A parametric study of prefrontal cortex involvement in human working memory. *Neuroimage* 5, 49–62.

- Bull, A.R., Cuddy, L.L., 1972. Recognition memory for pitch of fixed and roving stimulus tones. *Percept. Psychophys.* 11, 105–111.
- Carlson, S., Martinkauppi, S., Rämä, P., Salli, E., Korvenoja, A., Aronen, H.J., 1998. Distribution of cortical activation during visuospatial n-back tasks as revealed by functional magnetic resonance imaging. *Cereb. Cortex* 8, 743–752.
- Casey, B.J., Cohen, J.D., Jezzard, P., Turner, R., Noll, D.C., Trainor, R.J., Giedd, J., Kaysen, D., Hertz-Pannier, L., Rapoport, J.L., 1995. Activation of prefrontal cortex in children during a nonspatial working memory task with functional MRI. *Neuroimage* 8, 221–229.
- Clément, S., Demany, L., Semal, C., 1999. Memory for pitch versus memory for loudness. *J. Acoust. Soc. Am.* 106, 2805–2811.
- Cowan, N., Elliott, E.M., Saults, J.S., Morey, C.C., Mattox, S., Hismjatullina, A., Conway, A.R.A., 2005. On the capacity of attention: its estimation and role in working memory and cognitive aptitudes. *Cognit. Psychol.* 51, 42–100.
- Darwin, C.J., Hukin, R.W., 1999. Auditory objects of attention: the role of interaural time differences. *J. Exp. Psychol. Hum. Percept. Perform.* 25, 617–629.
- Demany, L., Montandon, G., Semal, C., 2004. Pitch perception and retention: two cumulative benefits of selective attention. *Percept. Psychophys.* 66, 609–617.
- Deutsch, D., 1970a. Dislocation of tones in a musical sequence: a memory illusion. *Nature* 226, 286–288.
- Deutsch, D., 1970b. Tones and numbers: specificity of interference in immediate memory. *Science* 168, 1604–1605.
- Deutsch, D., 1972a. Mapping of interactions in the pitch memory store. *Science* 175, 1020–1022.
- Deutsch, D., 1972b. Effect of repetition of standard and comparison tones on recognition memory for pitch. *J. Exp. Psychol.* 93, 156–162.
- Deutsch, D., 1974a. Generality of interference by tonal stimuli in recognition memory for pitch. *Quart. J. Exp. Psychol.* 26, 229–234.
- Deutsch, D., 1974b. An auditory illusion. *Nature* 251, 307–308.
- Deutsch, D., 1975. Facilitation by repetition in recognition memory for tonal pitch. *Mem. Cognit.* 3, 263–266.
- Deutsch, D., 1978a. Interference in pitch memory as a function of ear input. *Quart. J. Exp. Psychol.* 30, 283–287.
- Deutsch, D., 1978b. Lateralization by frequency for repeating sequences of dichotic 400 Hz and 800 Hz tones. *J. Acoust. Soc. Am.* 63, 184–186.
- Deutsch, D., 1980. Ear dominance and sequential interactions. *J. Acoust. Soc. Am.* 67, 220–228.
- Deutsch, D., 1988. Lateralization and sequential relationships in the octave illusion. *J. Acoust. Soc. Am.* 83, 365–368.
- Deutsch, D., 1999. The processing of pitch combinations. In: Deutsch, D. (Ed.), *The Psychology of Music*, second ed. Academic Press, San Diego, pp. 349–412.
- Deutsch, D., 2004. The octave illusion revisited again. *J. Exp. Psychol. Hum. Percept. Perform.* 30, 355–364.
- Deutsch, D., Feroe, J., 1975. Disinhibition in pitch memory. *Percept. Psychophys.* 17, 320–324.
- Deutsch, D., Roll, P.L., 1976. Separate “what” and “where” decision mechanisms in processing a dichotic tonal sequence. *J. Exp. Psychol. Hum. Percept. Perform.* 2, 23–29.
- Elliot, L.L., 1970. Pitch memory for short tones. *Percept. Psychophys.* 8, 379–384.
- Engle, R.W., Tuholoski, S.W., Laughlin, J.E., Conway, A.R.A., 1999. Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *J. Exp. Psychol.* 128, 309–331.
- Harris, J.D., 1952. The decline of pitch discrimination with time. *J. Exp. Psychol.* 43, 96–99.
- Kallman, H.J., Cameron, P.A., Beckstead, J.W., Joyce, E., 1987. Ear of input as a determinant of pitch-memory interference. *Mem. Cognit.* 15, 454–460.
- Koester, T., 1945. The time error in pitch and loudness discrimination as a function of time interval and stimulus level. *Arch. Psychol.* 297.
- Kuhn, G.F., 1977. Model for the interaural time differences in the azimuthal plane. *J. Acoust. Soc. Am.* 82, 157–167.
- Maeder, P.P., Meuli, R.A., Andriai, M., Bellman, A., Fornari, E., Thiran, J.P., Pittet, A., Clarke, S., 2001. Distinct pathway involved in sound recognition and localization: a human fMRI study. *Neuroimage* 14, 802–816.
- Martinkauppi, S., Rämä, P., Aronen, H.J., Korvenoja, A., Carlson, S., 2000. Working memory of auditory localization. *Cereb. Cortex* 10, 889–898.
- Massaro, D.W., 1970. Retroactive interference in short-term recognition memory for pitch. *J. Exp. Psychol.* 83, 32–39.
- Pechmann, T., Mohr, G., 1992. Interference in memory for tonal pitch: implications for a working memory model. *Mem. Cognit.* 20, 314–320.
- Rauschecker, J.P., Tian, B., 2000. Mechanisms and streams for processing “what” and “where” in auditory cortex. *Proc. Natl. Acad. Sci. USA* 97, 11800–11806.
- Ryan, T.A., 1959. Comments on orthogonal components. *Psychol. Bull.* 56, 394–396.
- Semal, C., Demany, L., 1991. Dissociation of pitch from timbre in auditory short-term memory. *J. Acoust. Soc. Am.* 89, 2404–2410.
- Semal, C., Demany, L., 1993. Further evidence for an autonomous processing of pitch in auditory short-term memory. *J. Acoust. Soc. Am.* 94, 1315–1322.
- Vuontela, V., Steenari, M.R., Carlson, S., Koivisto, J., Fjällberg, M., Aronen, E.T., 2003. Audiospatial and visuospatial working memory in 6–13 year old school children. *Learn. Mem.* 10, 74–81.
- Wickelgren, W.A., 1966. Consolidation and retroactive interference in short-term recognition memory for pitch. *J. Exp. Psychol.* 72, 250–259.
- Wickelgren, W.A., 1969. Associative strength theory of recognition memory for pitch. *J. Math. Psychol.* 6, 13–61.
- Zwicker, T., 1984. Experimente zur dichotischen Oktav Tauschung [Experiments on the dichotic octave illusion]. *Acoustica* 55, 128–136.