Individual differences in the sensitivity to pitch direction

Catherine Semal and Laurent Demany[a)]
Laboratoire de Neurophysiologie, UMR CNRS 5543, BP 63, Université Victor Segalen, 146 rue Leo Saignat, F-33076 Bordeaux, France

(Received 10 February 2006; revised 16 August 2006; accepted 31 August 2006)

It is commonly assumed that one can always assign a direction—upward or downward—to a percept of pitch change. The present study shows that this is true for some, but not all, listeners. Frequency difference limens (FDLs, in cents) for pure tones roved in frequency were measured in two conditions. In one condition, the task was to *detect* frequency changes; in the other condition, the task was to *identify* the direction of frequency changes. For three listeners, the identification FDL was about 1.5 times smaller than the detection FDL, as predicted (counterintuitively) by signal detection theory under the assumption that performance in the two conditions was limited by one and the same internal noise. For three other listeners, however, the identification FDL was much larger than the detection FDL. The latter listeners had relatively high detection FDLs. They had no difficulty in identifying the direction of just-detectable changes in intensity, or in the frequency of amplitude modulation. Their difficulty in perceiving the direction of small frequency/pitch changes showed up not only when the task required absolute judgments of direction, but also when the directions of two successive frequency changes had to be judged as identical or different. © *2006 Acoustical Society of America.* [DOI: 10.1121/1.2357708]

PACS number(s): 43.66.Hg, 43.66.Fe [AJO] Pages: 3907–3915

I. INTRODUCTION

The American National Standards Institute (ANSI, 1994) defines pitch as "that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high." This definition tallies with the fact that, in numerous languages, the expression meaning "pitch" (e.g., *hauteur tonale* in French, *Tonhöhe* in German, *visina tona* in Croatian and Serbian) incorporates a notion of height. Several authors (e.g., Bachem, 1950) have argued that pitch is not reducible to "tone height" because two tones one octave apart, which are quite distinct with respect to tone height, share at the same time another pitch quality ("tone chroma"). Nevertheless, it is generally believed that one can always assign a *direction*—upward or downward—to a percept of pitch change. Is this actually true for every listener, regardless of the size of the pitch change?

Using pure-tone stimuli, Jesteadt and Bilger (1974) measured the frequency discrimination abilities of four listeners (ordinary students, apparently) in several psychophysical paradigms. One of these paradigms was a two-interval forced-choice (2IFC) task; on each trial, two successive tones differing in frequency (by ΔF) were presented and the listener had to indicate whether the second tone was higher or lower than the first; the frequency of the first tone varied randomly from trial to trial ("roving" procedure), between 795 and 1260 Hz. In a second paradigm, two successive tones were again presented on each trial, but this time they could be either identical or different in frequency and the listener had to make a same/different (SD) judgment; when the two tones differed from each other, the second tone was always higher; the frequency of the first tone was again roved from trial to trial. Whereas the 2IFC task required a sensitivity to the direction of frequency changes (and thus pitch changes, presumably), this was not the case for the SD task. For each task, Jesteadt and Bilger measured the slope of the individual psychometric functions ($d'/\Delta F$). Under the assumption that for both tasks the decision variable was a *signed* pitch difference, signal detection theory (Green and Swets, 1974) predicted that the slopes of the psychometric functions would be two times higher in the 2IFC task than in the SD task. The experimental results appeared to be consistent with this prediction. Therefore, Jesteadt and Bilger's study suggests that as soon as a pitch change between two tones is detected, its direction can be identified.

Contrary to Jesteadt and Bilger, however, Wickelgren (1969) suggested that SD judgments in the frequency/pitch domain are not based on the same internal variable as are higher/lower judgments. In Wickelgren's study, three listeners were presented with sequences of three tones (T1, T2, T3). On each trial, the frequency of T1 was selected randomly between 400 and 490 Hz, and T3 could be higher, lower, or identical to T1. The task was to identify the relation between T3 and T1 using three response categories (higher, lower, same) and a three-level confidence rating; T2 had a fixed frequency, always remote from those of T1 and T3. An analysis of the receiver operating characteristics (ROCs; cf. Green and Swets, 1974, chap. 2) led Wickelgren to argue that listeners' judgments were partly based on the relative "familiarity" of the pitch evoked by T3, a variable depending on the *unsigned* difference between T1 and T3. Wickelgren posited that this variable of familiarity is the main determinant of SD judgments on pitch when the differences to be detected are small.

[a)]Author to whom correspondence should be addressed. Telephone: +33 55757 1651; Fax: +33 55690 1421.
Electronic mail: laurent.demany@psyac.u-bordeaux2.fr

More recently, Neuhoff *et al.* (2002) performed on a large group of students (tested in an auditorium) a short experiment which was similar to that of Wickelgren, but in which there was no tonal interference between the two tones to be compared on each trial. Neuhoff *et al.* found that when the two tones differed in frequency, a large majority of errors consisted of incorrect judgments of direction ("higher" instead of "lower," or vice versa) rather than "same" responses. This finding was consistent with Wickelgren's main point. One possible interpretation of the results was that, for a significant fraction of human listeners, upward and downward pitch changes are *perceptually* confusable. However, Neuhoff *et al.* did not interpret their data in this way. They suggested instead that the incorrect judgments of direction were due partly to a failure to detect some shifts and partly, in the authors' words, to "conceptual errors," that is, the "lack of knowledge of the appropriate labels for rising and falling pitch."

In a fourth study, by Johnsrude *et al.* (2000), two groups of listeners were tested. One group was neurologically normal and the other consisted of patients with lesions in the left or right temporal lobe. In both groups, using adaptive procedures, the authors measured frequency difference limens (FDLs) for pure tones close to 1000 Hz. As in Jesteadt and Bilger's (1974) study, the listeners were required to perform, in separate blocks of trials, an SD task and a 2IFC task (higher/lower judgments). The mean FDLs measured in the SD task were similar for the two groups. In the 2IFC task, the mean FDLs of patients with lesions in the left temporal lobe were also similar to those of the normal group, but performance was significantly poorer for patients with lesions in the right temporal lobe when these lesions affected the gyrus of Heschl. This demonstrated a dissociation between the ability to detect frequency changes and the ability to identify the direction of such changes. Johnsrude *et al.* concluded that: (1) The ability to identify pitch direction crucially depends on neural processes taking place in the gyrus of Heschl of the right hemisphere; (2) lower-order neural processes are sufficient for pitch change detection. It should be noted, however, that the mean FDLs measured in normal listeners by Johnsrude *et al.* were very much higher than those previously obtained by other authors in very similar experimental conditions (see, e.g., Sek and Moore, 1995). The subjects of Johnsrude *et al.* were apparently not trained and it is likely that, after some training, their performances would have been quite different (see, e.g., Demany and Semal, 2002). In another neuropsychological investigation which was methodologically similar to that of Johnsrude *et al.*, Tramo *et al.* (2002) tested in particular a patient with bilateral lesions of the auditory cortex. This patient appeared to have abnormally high FDLs in the SD task as well as the 2IFC task. However, his deficit relative to normal controls was definitely larger in the 2IFC task, in accordance with the dissociation reported by Johnsrude *et al.*

The foregoing literature review shows that conflicting findings and conclusions have been reported regarding human listeners' sensitivity to pitch direction in small frequency/pitch changes. The starting point of the present research was the informal observation, by the authors, of a pronounced difference between them in this respect. Both authors are neurologically and audiologically normal. Moreover, both are able to detect small frequency changes in the absence of loudness cues and, therefore, neither of them is "tone deaf" (Peretz and Hyde, 2003; Foxton *et al.*, 2004). However, one of the authors is completely unable to identify the direction of small, but perfectly detectable, frequency/pitch changes, whereas this problem does not exist for the other author. The experiments reported here were intended to confirm the reality of such individual differences.

II. EXPERIMENT 1

In this experiment, performed on nine normal-hearing listeners, we measured FDLs in two conditions assessing, respectively, the ability to detect frequency changes and the ability to identify their direction. The frequencies and sound pressure level (SPL) of the pure tones used as stimuli were such that the perceptual correlate of a frequency change could always be assumed to be nothing but a pitch change. Trials had exactly the same form in the two conditions, and a two-alternative forced choice task was used in each case. The two conditions differed from each other only with respect to the question asked on each trial. The observed relations between the two types of FDLs were compared to the relation predicted by signal detection theory for an ideal listener.

A. Method

1. Subjects

The nine subjects (L1, L2, L3, …L9) included the authors (L1 and L7) and seven college students who were paid for their services. Three of the students (L4, L8, and L9) were selected among 68 persons who were pre-tested collectively in an auditorium. The aim of this pre-test was to identify and recruit listeners having difficulties regarding the perception of pitch direction.[1] All subjects had normal audiograms from 125 to 4000 Hz, and considered themselves as neurologically healthy. None of the subjects had a thorough musical education, but two of them (L1 and L3) were amateur musicians. L1 and L7 had a considerable prior experience in psychoacoustical tasks (including frequency discrimination tasks), whereas the other subjects were novices.

2. Procedure

Subjects were tested individually in a sound-attenuating booth (Gisol, Bordeaux). On each trial, four successive pure tones were presented diotically, through headphones (Sennheiser HD265). They were generated via 24 bit digital-to-analog converters (RME), at a sampling rate of 44.1 kHz. They had equal amplitudes, a nominal SPL of 65 dB, a total duration of 250 ms, and were gated on and off with 20 ms cosinusoidal amplitude ramps. There was a silent interval (ISI) of 250 ms between the first two tones and between the last two tones. A longer ISI—700 ms—separated the second and third tones, thus segmenting the whole sequence into two pairs of tones. One of the pairs, selected at random, consisted of identical tones, whereas the two members of the other pair differed in frequency. The frequencies of the first members of the pairs were selected randomly, independently

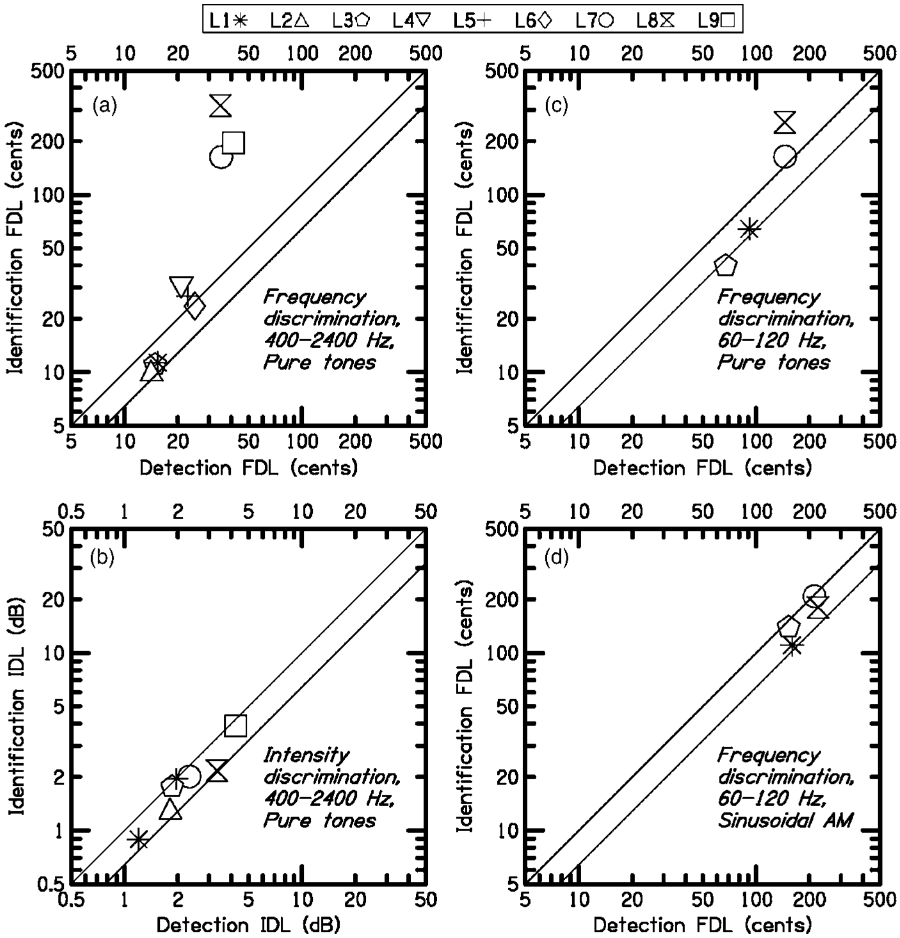

FIG. 1. (a) Detection and identification FDLs (frequency difference limens) of the nine listeners tested in experiment 1; FDLs were measured in musical cents and are here scaled logarithmically. (b) Detection and identification IDLs (intensity difference limens) of the seven listeners tested in experiment 4. (c) FDLs measured in experiment 5 for 60–120 Hz pure tones. (d) FDLs measured in experiment 5 for 60–120 Hz sinusoidal amplitude modulations. In each panel, the lower oblique line represents a prediction of signal detection theory (see the text).

of each other, between 400 and 2400 Hz (the probability distribution being rectangular on a log-frequency scale). In the pair including different tones, frequency changed either upwards or downwards, equiprobably.

FDLs were measured in a "detection" condition and an "identification" condition, corresponding to separate blocks of trials. In the detection condition, the task was to indicate on each trial if the pair including different tones was the first or the second pair. In the identification condition, the task was to identify the direction (upward or downward) of the frequency change that occurred within one pair, without specifying if this was the first or second pair. Responses were given by making a mouse-click on one of two labeled zones of a monitor screen, and were immediately followed by visual feedback. Response times were unlimited. Within a block of trials, there was a 600 ms pause between each response and the first tone of the next trial.

The critical frequency change occurring on each trial (ΔF) was defined in musical cents rather than in Hertz (1 cent=1/100 semitone=1/1200 octave). It was desirable to do so because, from 400 to 2400 Hz, FDLs vary markedly in terms of Hertz but are roughly constant in relative terms (see, e.g., Sek and Moore, 1995). In all blocks of trials, $|\Delta F|$ was initially large. It was decreased following each correct response, and increased following each incorrect response. A block ended after the 14th reversal in the variation of $|\Delta F|$. Up to the fourth reversal, $|\Delta F|$ was multiplied by 2.25 when it was increased, and divided by the cube root of the same factor when it was decreased. After the fourth reversal, $|\Delta F|$ was either multiplied by 1.5 or divided by the cube root of this factor. The FDL measured in a block of trials was defined as the geometric mean of all the $|\Delta F|$ values used from the fifth reversal on. This statistic estimated the 75% correct point of the psychometric function (Kaernbach, 1991).

In each test session, FDLs were measured alternately in the detection condition and the identification condition. The switch occurred after each FDL measurement. The number of FDL measurements varied from session to session but overall, for each listener, 50 FDL measurements were made in each condition; this corresponded to approximately 2000 trials per condition.

B. Results and discussion

For each condition and listener, we computed the geometric mean of the 50 measured FDLs. The results are displayed in Fig. 1(a). Here, each listener is represented by a specific symbol, as indicated on the top of the figure. The horizontal and vertical coordinates represent detection and identification performances, respectively, on identical logarithmic scales.

Two parallel oblique lines are drawn in each panel. The upper oblique line—i.e., the diagonal—indicates where the data points should be if identification performance [FDLs in the case of Fig. 1(a)] was equal to detection performance. Under the assumptions of "high-threshold theory" (Green and Swets, 1974, chap. 5), if the detection of a frequency/pitch change was always sufficient for the identification of its direction, then the data points should be on this diagonal; moreover, it should be impossible to find a data point significantly below the diagonal because this would mean that the direction of *non*detectable frequency changes can be identified. The lower oblique line, on the other hand, represents the expected loci of the data points for an ideal listener defined in the framework of signal detection theory. For the ideal listener (Λ) in question: (1) A given pitch sensation is a point on an internal low-high continuum that is mathematically equivalent to a logarithmically scaled frequency axis; (2) performance in the two experimental conditions is limited only by a Gaussian random variability of the pitch sensations evoked by a tone of fixed frequency; (3) the variance of pitch for a given frequency corresponds to a given frequency ratio, independent of frequency. The strategy of Λ on each trial is optimal. In the detection condition, therefore, Λ measures the size of the pitch change elicited by each pair of tones, and then votes for the pair for which the change is larger. In the identification condition, Λ votes for the direction of the larger of the two pitch changes. We determined by means of computer simulations that in order to obtain, from Λ, 75% of correct responses in the two conditions, the signal (i.e., the frequency change presented in one pair of tones) had to be 1.56 times larger in the detection condition than in the identification condition.[2] This factor of 1.56 corresponds to the vertical or horizontal distance of the two oblique lines in each panel of Fig. 1.

Consider now the nine data points of Fig. 1(a). In the detection condition, the geometric standard errors of the nine mean FDLs had a mean value of 4.8% and ranged from 3.6% to 7.6%. In the identification condition, the corresponding standard errors had a mean value of 8.0% and ranged from 4.4% to 12.8%. So, the precision of the data is such that, vertically as well as horizontally, each of the nine symbols displayed in Fig. 1(a) has a size exceeding two standard errors. It can be seen that the nine data points form three clusters of three points. One cluster includes the data of three listeners (L1, L2, L3) for whom the detection FDLs were lowest (around 15 cents). For each of these three listeners, the identification FDL was lower than the detection FDL, and in each case the corresponding trend was statistically significant $[t(98) \geq 3.93;\ P < 0.001]$. The three data points are close to the lower oblique line, which means that L1, L2, and L3 behaved almost exactly like the ideal listener Λ defined above.[3] Let us point out that even for this subgroup of "best" listeners, the obtained detection and identification FDLs were higher than the FDLs reported by a number of authors (e.g., Wier *et al.*, 1977; Nelson *et al.*, 1983; Sek and Moore, 1995). This may be largely due to the fact that FDLs are generally measured for fixed standard tones, whereas we used here a roving procedure (see in this respect Demany and Semal, 2005).

L4, L5, and L6 produced a second cluster. Their detection FDLs (around 20 cents) are somewhat poorer that those of L1, L2, and L3. Their identification FDLs are approximately equal to their detection FDLs.

L7, L8, and L9 produced the third cluster. Their detection FDLs (around 35 cents) were the poorest. In the detection condition, however, these listeners' FDLs were only two or three times larger than those of L1, L2, or L3. By contrast, their identification FDLs were dramatically poor: They range from 163 to 317 cents. The ratio of the two types of FDL (identification/detection) is equal to 4.7 for L7, 4.8 for L9, and as much as 9.2 for L8. It is clear that these three listeners were completely unable to identify the direction of frequency changes that they nonetheless detected perfectly.

Globally, there was a correlation (Pearson r) of 0.90 between the logarithms of the detection FDLs and the logarithms of the ratios of the two types of FDL. This correlation is statistically significant ($\mathrm{d.f.}=7$, $P < 0.01$).

In the experiment, as mentioned above, the response given on each trial was followed by visual feedback. In the identification condition, the feedback should have solved very rapidly the problem of a subject who would be able to perceive pitch direction without any difficulty but would not know the appropriate response for each direction. However, the feedback was also liable to have a more protracted benefit for listeners having difficulty in perceiving pitch direction: Thanks to the feedback, such listeners might have progressively learned to perceive pitch direction. This hypothesis led us to analyze the evolution of the FDLs measured in L7, L8, and L9 during the experiment. For each condition and listener, the geometric means of the FDLs measured in trial blocks $1-5, 6-10, \ldots 46-50$ are displayed in Fig. 2, where open and closed symbols represent detection and identification FDLs, respectively. The identification FDLs of L9 did improve during the experiment, and got closer and closer to her detection FDLs. For L7 and L8, in contrast, there was no sign of perceptual learning.

III. EXPERIMENT 2

In experiment 1, L7, L8, and L9 were unable to identify the direction of *frequency* changes that they nonetheless detected perfectly. This suggests that L7, L8, and L9 can perceive a *pitch* change without being able to identify its direction. However, an alternative interpretation of these listeners' results in experiment 1 was possible. It could be argued that they detected ΔF on the basis of some cue other than pitch, and that it was only for $|\Delta F|$ values close to the identification FDLs that they began to perceive ΔF as a genuine change in pitch. If so, all the listeners tested in experiment 1 would in fact be able to identify the direction of any pitch change detectable as such.

What could be the "cue other than pitch" used by L7, L8, and L9 in the detection condition of experiment 1? Since these listeners had normal audiograms, it is very unlikely that they detected ΔF on the basis of loudness changes. Another conceivable hypothesis is that, when they were presented with a given pair of tones, they monitored the excitation level of a single auditory filter activated by both tones. A change in frequency from the first tone to the second tone produced a change in the excitation level of the filter, thus allowing the listeners to perform not too badly in the detection condition. On the other hand, if the filter was chosen at random, on either side of the tones' excitation patterns, and if

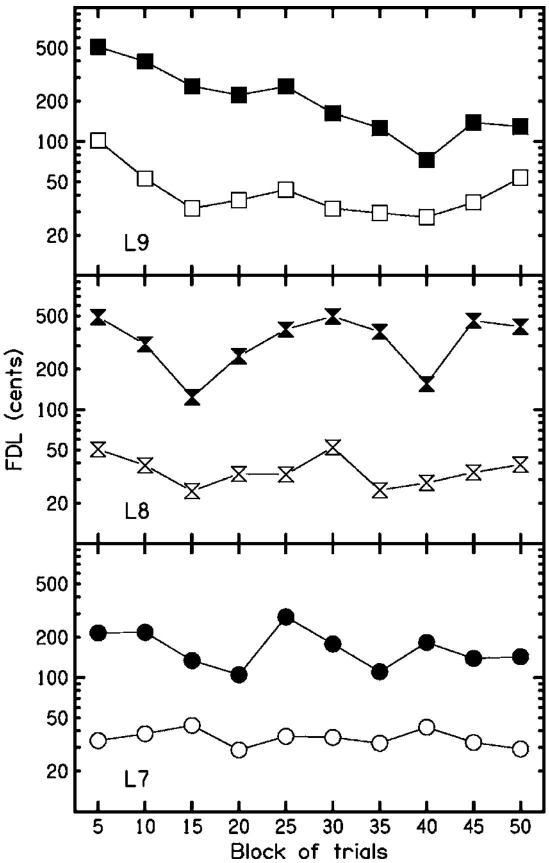

FIG. 2. Evolution of the FDLs measured in listeners L7, L8, and L9 during experiment 1. Open and closed symbols respectively represent detection and identification FDLs. Each data point is the geometric mean of five consecutive FDL measurements.

its center frequency was ignored, the direction of the change in excitation level did not provide information about the direction of the frequency change, which could explain the difficulty of the identification condition.

In experiment 2, this single-filter hypothesis was tested on L7 and L8 by replicating the detection condition of experiment 1 with the addition of random changes in intensity from tone to tone. In half of the FDL measurements, i.e., 20 blocks of trials for each subject, the SPL of every tone could take equiprobably any value between 60 and 70 dB. In the other half, the SPL was fixed at 65 dB, as in experiment 1. These two types of blocks were run alternately. If, on any given trial, L7 and L8 were monitoring the excitation level of a single auditory filter, then their FDLs should have been markedly larger with than without the intensity roving.

This was not the case. For L7, the geometric means of the FDLs obtained with and without the intensity roving were, respectively, 24.4 and 24.6 cents. For L8, the corresponding figures were 44.7 and 32.5 cents. So, on average, the roving of intensity increased the FDL by merely 17%. Moore and Glasberg (1989) and Emmerich *et al.* (1989) performed experiments comparable to the present one on listeners who had no difficulty in identifying pitch direction. The intensity roving range used by Moore and Glasberg (6 dB) was smaller than ours (10 dB). In their experiment, nonetheless, intensity roving had an average effect of the same size

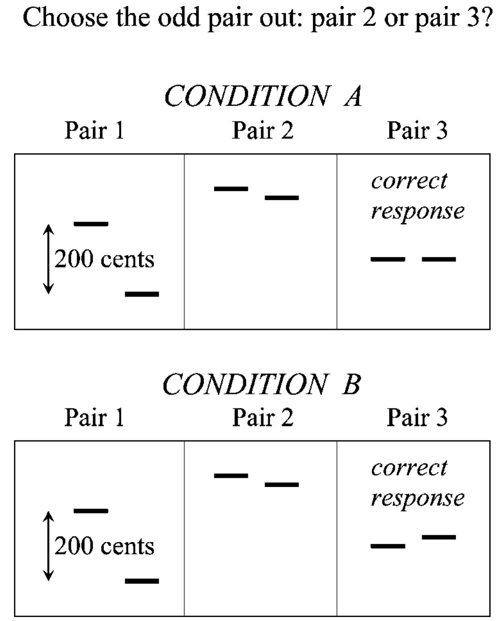

FIG. 3. The two conditions of experiment 3, illustrated by an example.

as the effect observed here. Emmerich *et al.* (1989), who used a 20 dB roving range, obtained a much larger effect.

In conclusion, experiment 2 did not support the idea that, in the detection condition of experiment 1, L7 and L8 used a perceptual cue other than pitch. It appears in any case that these listeners have no difficulty in dissociating pitch from loudness.

IV. EXPERIMENT 3

In the identification condition of experiment 1, subjects had to make *absolute* judgments on the direction of frequency changes. It is conceivable that a listener having difficulty in this task would nonetheless be able to perceive that a frequency change in a given direction is more similar to another frequency change in the same direction than to a frequency change in the opposite direction. The main goal of experiment 3 was to determine if this was true for L7, L8, and L9, the three listeners who had very poor identification FDLs in experiment 1.

A. Method

L1, L2, L3, L7, L8, and L9 were tested in two conditions, illustrated in Fig. 3. On each trial, three successive pairs of pure tones were presented. Within two of the three pairs, frequency changed in the same direction. In the remaining pair, there was either no frequency change (condition A) or a frequency change in the opposite direction (condition B). This "odd" pair was either Pair 2 or Pair 3, at random, and the subject's task was to identify its position in a 2AFC paradigm. The two elements of Pair 1 were always 200 cents apart, but the direction of the corresponding frequency change varied randomly from trial to trial. In condition B, therefore, the direction of the frequency change made in the odd pair also varied randomly from trial to trial. However, on every trial run in condition B, the changes made in Pairs 2 and 3 had identical sizes (in cents). As in experiment

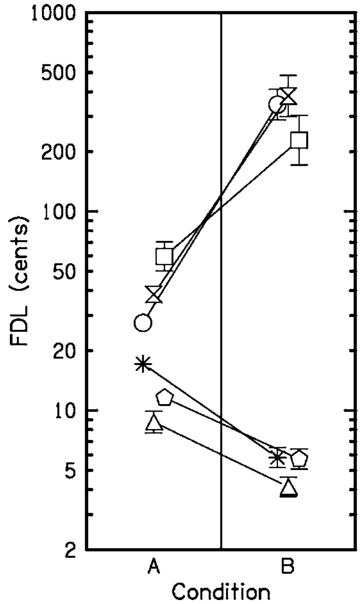

FIG. 4. Results of experiment 3. Each listener is represented by a specific symbol, the same as in Fig. 1. The error bars represent geometric standard errors. Error bars smaller than the symbol indicating the mean are not visible.

1, the tones had equal amplitudes, a nominal sound pressure level of 65 dB, and a duration of 250 ms. The ISI was again 250 ms within pairs and 700 ms between pairs. The frequency of the first member of each pair was again selected randomly between 400 and 2400 Hz. Visual feedback was once more provided following each response.

In each condition, the size of the frequency change ΔF made in Pair 2 and/or Pair 3 was varied from trial to trial, according to the same adaptive procedure as that used in experiment 1. We thus measured FDLs corresponding to the value of $|\Delta F|$ for which P(C) was 75%. For L1, L2, and L3, who can identify the direction of frequency in just-detectable frequency changes, our prediction was that FDLs would be lower in condition B than in condition A; this was to be expected because, for a given value of $|\Delta F|$, the difference between Pairs 2 and 3 was larger in condition B than in condition A. On the other hand, an opposite prediction was made for L7, L8, and L9, who were unable in experiment 1 to identify the direction of frequency changes well above their detection threshold; for these listeners, it could be expected that condition B would be the more difficult condition because it crucially required perceptual sensitivity to the direction of frequency changes, whereas this was not the case in condition A.

Within each test session, as in experiment 1, FDLs were measured alternately in the two conditions. The total number of FDL measurements per condition was equal to 6 for L1, 25 for L2, 10 for L3, 15 for L7 and L8, and 18 for L9.

B. Results and discussion

Figure 4 displays the geometric mean of the FDL measurements made for each condition and listener, as well as the associated standard errors. The results are extremely clear-cut and they agree with the predictions stated above: The FDLs of L1, L2, and L3 were lower in condition B than in condition A, whereas the opposite was true for L7, L8, and L9. In both conditions, L7, L8, and L9 were less efficient than L1, L2, and L3. However, the two groups do not differ dramatically from each other in condition A: Their average FDLs are in a ratio of about 3 for this condition. In condition B, by contrast, their average FDLs differ by a factor of about 60. Let us note that, in the blocks of trials estimating the FDL of L7, L8, and L9 in condition B, the initial value of $|\Delta F|$ was generally smaller than the FDL estimate eventually obtained. In consequence, the true FDLs of these listeners in condition B may be even higher than the estimates displayed in Fig. 4.

It is not surprising that L7, L8, and L9 were so inefficient in condition B given that the frequency change occurring in Pair 1 had a magnitude of 200 cents. This magnitude of change was well above all the detection FDLs measured in experiment 1, but not larger than the mean identification FDL of L7, L8, and L9 (216 cents). These three listeners might be somewhat more efficient in condition B if the frequency change produced in Pair 1 were larger. In any case, the present data make clear that their poor performance in the identification condition of experiment 1 is not due to the fact that the task required absolute judgments: L7, L8, and L9 show a perceptual insensitivity to the direction of frequency changes in relative as well as absolute judgments.

V. EXPERIMENT 4

We have shown above that there are pronounced individual differences regarding the perception of direction in frequency—and presumably pitch—changes. Is pitch special from that point of view? This question led us to perform a variant of experiment 1 in which the acoustic changes to be detected and identified as upward or downward changes were changes in intensity rather than frequency. Except for this novelty, the procedure and stimuli were identical to those of experiment 1. On each trial, therefore, four tones were presented in two pairs, and one tone (tone 2 or tone 4) differed in intensity from the other three (all at 65 dB). The intensity difference, ΔI, was at random positive or negative. Its absolute value $|\Delta I|$ was varied adaptively from trial to trial in order to find intensity difference limens (IDLs); the dB value of $|\Delta I|$ was manipulated exactly like $|\Delta F|$ (in cents) previously. On a given trial, the two elements of each pair of tones now had the same frequency. However, the frequencies of tones 1 and 3 were, as before, selected at random and independently of each other; in consequence, there was again a generally large frequency change from the first pair to the second pair. Since intensity comparisons between pure tones are markedly easier and more accurate for tones of the same frequency than for tones with different frequencies (Lim *et al.*, 1977), it was reasonable to assume that the IDLs measured here would be based exclusively on within-pair comparisons, as was the case for the FDLs measured in experiment 1.

The tested listeners included all those previously tested in experiment 1, except for L4 and L6. For each listener (with the exception of L9), 50 IDL measurements (15 for L9) were made in the detection condition and in the identification condition. The results are displayed in Fig. 1(b). Each of the seven data points lies between the two oblique lines. Thus, the performance of all listeners was consistent with the idea that they could identify the direction of an intensity change as soon as they could detect it. In this respect, L7, L8, and L9 did not behave at all here as in experiment 1. Note, however, that they were in both experiments the three subjects for whom performance was poorest in the detection condition.

VI. EXPERIMENT 5

It could still be hypothesized, after experiment 4, that the perceptual problem of listeners such as L7, L8, and L9 is not exclusively a difficulty to perceive the direction of *pitch*, but more generally a difficulty to perceive the direction of *frequency*. The frequency of, for instance, a sinusoidal amplitude modulation imposed on a pure tone does not evoke a salient pitch percept when this modulation frequency is at least 15 times lower than the carrier frequency (Ritsma, 1962). On the other hand, a (sufficiently large) change in the modulation frequency can be heard (by many listeners, at least) as an increase or decrease in the number of modulation cycles per unit of time; by contrast, pure tones are perceived as smooth. In experiment 5, we replicated the procedure of experiment 1 on: (1) A continuum of *modulation* frequency, ranging from 60 to 120 Hz (for the first tone in each pair); (2) an *audio*-frequency continuum also ranging from 60 to 120 Hz. In the first of these two conditions, the modulations were sinusoidal and imposed on a sinusoidal carrier of 2000 Hz and 61 dB SPL. In the second condition, the stimuli were pure tones, approximately equalized in loudness at a level of 65 phons (by decreasing the SPL continuously from 85 to 75 dB between 60 and 120 Hz). The duration of all stimuli was set to 500 ms.

The experiment was performed on L1, L3, L7, and L8. For each continuum and listener, 15 FDL measurements were made in the detection condition and the identification condition. The results are displayed in Figs. 1(c) and 1(d). Not surprisingly, all the detection FDLs were much higher than those obtained in experiment 1. When the stimuli were pure tones [Fig. 1(c)], L1 and L3 behaved once more almost exactly like the ideal listener Λ, and for L8 the identification FDL was once more definitely larger than the detection FDL. However, contrary to what had been found in experiment 1, the identification FDL of L7 was very close to her detection FDL. When the manipulated frequency was a modulation frequency [Fig. 1(d)], the four listeners behaved similarly and none of them appeared to encounter a difficulty in the identification condition. It is remarkable that for L8, from Fig. 1(c) to Fig. 1(d), the detection FDL increases whereas the identification FDL decreases. Both of these trends are statistically significant [$t(28) \geqslant 2.53$; $P \leqslant 0.017$]. It is clear, therefore, that the perceptual problem of this listener is a difficulty to identify the direction of *pitch* rather than frequency per se.

For both L7 and L8, when the stimuli were pure tones, identification performance was not very much poorer than detection performance, in contrast to what had been found in experiment 1. Did this reflect an improvement, due to learning, in the ability of these listeners to perceive pitch direction? To answer that question, we retested L7 and L8 in the detection and identification conditions of experiment 1. A total of 20 new FDL measurements were made for each listener. The obtained geometric means of the detection and identification FDLs were, respectively, 24.8 and 317.8 cents for L7, and 37.3 and 238.8 cents for L8. Thus, the learning hypothesis had to be rejected. We believe instead that, for these listeners, a pitch change must exceed an approximately fixed magnitude in order to be identifiable as an upward change or a downward change. This conjecture makes sense of the fact that the identification FDLs of L7 and L8 for pure tones had similar values in experiments 1 and 5 (see Fig. 1).

VII. GENERAL DISCUSSION

As pointed out in the Introduction, Johnsrude *et al.* (2000) argued that the detection of frequency/pitch changes and the identification of their direction are distinct perceptual abilities, mediated by separate neural processes. The behavior of three of our subjects (L7, L8, and L9) was consistent with this thesis since these three subjects proved unable to identify the direction of frequency changes that, nonetheless, they detected perfectly and apparently heard as pitch changes. Although the subjects in question were neurologically normal, they showed the same dissociation as that observed by Johnsrude *et al.* (2000) or Tramo *et al.* (2002) in certain patients with brain lesions. On the other hand, the behavior of three other participants in our study (L1, L2, and L3) was quite different: In experiment 1, their identification FDLs were significantly *lower* than their detection FDLs. Their data were approximately consistent with the predictions of a model assuming that the detection of a frequency change and the identification of its direction are limited by one and the same internal noise: a random variability of the pitch sensations evoked by a tone of a given frequency.

In experiment 1, we also found that the ability to identify frequency direction in a just-detectable frequency change was strongly correlated with the detection FDL itself: The subjects who were efficient in the identification task were also efficient in the detection task. This result is remarkable because it suggests that, contrary to the thesis of Johnsrude *et al.*, the neural processes underlying the detection of frequency/pitch changes and the identification of their direction are not independent of each other. One possible speculation is that there are two separate detection mechanisms: an optimal mechanism which is sensitive to the direction of changes, and a non-optimal mechanism which is not direction-sensitive.

Although L7, L8, and L9 were unable to identify the direction of one-semitone changes in frequency, they could identify the direction of larger changes, and were indeed relatively successful for changes amounting to a few semitones. This effect of size is difficult to understand. The listeners themselves have little to say about the nature of their perceptual problem. Not surprisingly, the three of them are unable to sing in tune; but they are able to identify well-known melodies, and L8 enjoys listening to music. The perceptual problem of some of the subjects who have difficulty in identifying the direction of small frequency/pitch changes may disappear following a long period of adequate training. However, this seems unlikely for L7 and L8 since, at the end of the present study, they had made no progress after more than 2000 identification trials run using an adaptive procedure including feedback. It would be interesting to get some idea of the prevalence of their apparently irremediable perceptual difficulty in the general population; let us emphasize in this regard that the group tested here may not be a representative sample.

Paavilainen *et al.* (1999) reported electro-encephalographic data concerning the sensitivity of human listeners to the direction of frequency changes. Their subjects were presented with a long sequence of pure tone pairs with randomly varying frequencies. Each tone had a duration of 50 ms and there was a 40 ms ISI within pairs. Frequency changed upward in most (87.5%) of the pairs, and downward in the remaining pairs. The authors' aim was to determine if the infrequent downward changes would elicit a significant mismatch negativity (MMN). This was indeed the case for the tested group as a whole. Because the MMN is a largely automatic (attention-independent) brain response, the authors suggested that the human brain contains neuronal populations which are selectively sensitive to the direction of discrete frequency changes. In the experiment of Paavilainen *et al.*, it was apparently not attempted to find differences between listeners. However, given that the frequency changes used had a minimum size of two semitones and were generally much larger, a significant MMN might have been found even in our subjects L7, L8, and L9. It would be interesting to perform a similar experiment using smaller frequency changes and separate subgroups of listeners, with different behavioral abilities to identify pitch direction.

Demany and Ramos (2005) also supported the hypothesis that the human brain contains direction-sensitive detectors of discrete frequency/pitch changes. They did so on the basis of purely psychophysical experiments. On each trial, their subjects were presented with a random "chord" of five simultaneous pure tones, followed after an ISI by a single tone (T). Because the component tones of the chords were gated on and off synchronously, they were very hard to hear out individually. This was confirmed in a condition where, on any trial, T could be either identical to a randomly selected component of the chord or halfway in (log) frequency between two components: Discriminating between the two corresponding types of chord-T sequences appeared to be very difficult. However, in another condition where T was instead one-semitone higher or lower than a randomly selected component of the chord and the task was to vote for "higher" or "lower," performance was much better (even when the ISI was several seconds long). In the latter condition, it appeared that listeners were able to perceive consciously the direction of a pitch movement produced by two tones without perceiving consciously the pitch of the first tone. This result, and those of related experiments, seemed to constitute strong evidence for the existence of automatic and direction-sensitive "frequency-shift detectors" in the human brain. Demany and Ramos conjectured that the detectors in question respond to frequency shifts of less than one semitone, and indeed respond to any frequency shift which is large enough to be audible. If this is true, then the results of the present research imply that such detectors do not exist in the brain of some listeners. We are inclined to think instead that they exist in the brain of every normal listener but that, for some listeners, they do not respond to very small shifts.

ACKNOWLEDGMENTS

We thank Hedwig Gockel, Maja Serman, and especially Christophe Micheyl for judicious comments on a previous version of this paper.

[1]More than 15% of the 68 pre-tested listeners were definitely less efficient in the identification task than in the detection task. Within this group, we selected three listeners for whom the difference was particularly large.

[2]A qualitative explanation of this outcome goes as follows. Suppose that, on a given trial, the larger of the two pitch changes measured by Λ has been evoked by the pair of identical tones. In the detection condition, the response of Λ on such a trial will be wrong with a probability of 1. In the identification condition, by contrast, the probability of a wrong response is only 1/2, because the pitch change evoked by the pair of identical tones was equiprobably ascending or descending. On a trial for which the larger of the two pitch changes measured by Λ takes place in the "right" pair (i.e., the pair of different tones), this larger change can nonetheless be in the wrong direction. However, the probability of such an event is smaller than 1/2. So, globally, for a given $|\Delta F|$, Λ is expected to perform better in the identification condition than in the detection condition. The psychophysical paradigm used in our detection condition has been analyzed mathematically in several publications (Macmillan *et al.*, 1977; Rousseau and Ennis, 2001; Micheyl and Messing, in press). For this paradigm, Rousseau and Ennis (2001) [see also Micheyl and Oxenham (2005)] provided an equation relating d' (that is, the subject's sensitivity to the signal to be detected, in our case ΔF) to the probability of a correct response, P. This equation is: $d'=2\mathbf{\Phi}^{-1}\,[1/2+(P/2-1/4)^{1/2}]$, where $\mathbf{\Phi}^{-1}$ denotes the inverse of the cumulative standard normal. For the same signal ΔF, in our identification condition, it can be shown that, more simply: $d'=2\mathbf{\Phi}^{-1}(P)$ (Micheyl, Kaernbach, and Demany, in preparation). Thus, when $P=0.75$, d' is 1.56 times larger in the identification condition than in the detection condition. Under the assumption that d' is proportional to $|\Delta F|$ in cents, this implies that, for Λ, the detection FDL is 1.56 times higher than the identification FDL.

[3]Nonetheless, for L1 and L3, the ratio of the two types of FDL (detection/identification) was significantly smaller than 1.56, the ratio expected for an ideal listener Λ [$t(98)\geq 1.99$; $P\leq 0.049$, two-tailed tests]. For L2, there was a marginally significant trend in the same direction [$t(98)=1.87$; $P=0.065$].

ANSI (**1994**). American National Standard Acoustical Terminology. ANSI S1.1-1994 (ASA 111-1994), 34.

Bachem, A. (**1950**). "Tone height and tone chroma as two different pitch qualities," Acta Psychol. **7**, 80–88.

Demany, L., and Ramos, C. (**2005**). "On the binding of successive sounds: Perceiving shifts in nonperceived pitches," J. Acoust. Soc. Am. **117**, 833–841.

Demany, L., and Semal, C. (**2002**). "Learning to perceive pitch differences," J. Acoust. Soc. Am. **111**, 1377–1388.

Demany, L., and Semal, C. (**2005**). "The slow formation of a pitch percept beyond the ending time of a short tone burst," Percept. Psychophys. **67**, 1376–1383.

Emmerich, D. S., Ellermeier, W., and Butensky, B. (**1989**). "A reexamination of the frequency discrimination of random-amplitude tones, and a test of Henning's modified energy-detector model," J. Acoust. Soc. Am. **85**, 1653–1659.

Foxton, J. M., Dean, J. L., Gee, R., Peretz, I., and Griffiths, T. D. (**2004**). "Characterization of deficits in pitch perception underlying 'tone deafness'," Brain **127**, 801–810.

Green, D. M., and Swets, J. A. (**1974**). *Signal Detection Theory and Psychophysics* (Krieger, New York).

Jesteadt, W., and Bilger, R. C. (**1974**). "Intensity and frequency discrimination in one- and two-interval paradigms," J. Acoust. Soc. Am. **55**, 1266–1276.

Johnsrude, I. S., Penhune, V. B., and Zatorre, R. J. (**2000**). "Functional specificity in the right human auditory cortex for perceiving pitch direction," Brain **123**, 155–163.

Kaernbach, C. (**1991**). "Simple adaptive testing with the weighted up-down method," Percept. Psychophys. **49**, 227–229.

Lim, J. S., Rabinowitz, L. D., Braida, L. D., and Durlach, N. I. (**1977**). "Intensity perception. VIII. Loudness comparisons between different types of stimuli," J. Acoust. Soc. Am. **62**, 1577–1587.

Macmillan, N. A., Kaplan, H. L., and Creelman, C. D. (**1977**). "The psychophysics of categorical perception," Psychol. Rev. **84**, 452–471.

Micheyl, C., and Messing, D. P. (in press). "Likelihood ratio, optimum decision rules, and correct-response probabilities in the 4IAX paradigm," Percept. Psychophys.

Micheyl, C., and Oxenham, A. J. (**2005**). "Comparing *F*0 discrimination in sequential and simultaneous conditions," J. Acoust. Soc. Am. **118**, 41–44.

Moore, B. C. J., and Glasberg, B. R. (**1989**). "Mechanisms underlying the frequency discrimination of pulsed tones and the detection of frequency modulation," J. Acoust. Soc. Am. **86**, 1722–1732.

Nelson, D. A., Stanton, M. E., and Freyman, R. L. (**1983**). "A general equation describing frequency discrimination as a function of frequency and sensation level," J. Acoust. Soc. Am. **73**, 2117–2123.

Neuhoff, J. G., Knight, R., and Wayand, J. (**2002**). "Pitch change, sonification, and musical expertise: Which way is up?," *Proceedings of the 2002 International Conference on Auditory Display*, Kyoto, Japan.

Paavilainen, P., Jaramillo, M., Näätänen, R., and Winkler, I. (**1999**). "Neuronal populations in the human brain extracting invariant relationships from acoustic variance," Neurosci. Lett. **265**, 179–182.

Peretz, I., and Hyde, K. L. (**2003**). "What is specific to musical processing? Insights from congenital amusia," Trends in Cognitive Science **7**, 362–367.

Ritsma, R. J. (**1962**). "Existence region of the tonal residue," J. Acoust. Soc. Am. **34**, 1224–1229.

Rousseau, B., and Ennis, D. M. (**2001**). "A Thurstonian model for the dual pair (4IAX) discrimination method," Percept. Psychophys. **63**, 1083–1090.

Sek, A., and Moore, B. C. J. (**1995**). "Frequency discrimination as a function of frequency, measured in several ways," J. Acoust. Soc. Am. **97**, 2479–2486.

Tramo, M. J., Shah, G. D., and Braida, L. D. (**2002**). "Functional role of auditory cortex in frequency processing and pitch perception," J. Neurophysiol. **87**, 122–139.

Wickelgren, W. A. (**1969**). "Associative strength theory of recognition memory for pitch," J. Math. Psychol. **6**, 13–61.

Wier, C. C., Jesteadt, W., and Green, D. M. (**1977**). "Frequency discrimination as a function of frequency and sensation level," J. Acoust. Soc. Am. **61**, 178–184.