Continuous versus discrete frequency changes: Different detection mechanisms?

Laurent Demany[a)]
Laboratoire Mouvement, Adaptation, Cognition (UMR CNRS 5227), Université de Bordeaux,
BP 63, 146 Rue Leo Saignat, F-33076 Bordeaux, France

Robert P. Carlyon
MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, United Kingdom

Catherine Semal
Laboratoire Mouvement, Adaptation, Cognition (UMR CNRS 5227), Université de Bordeaux,
BP 63, 146 Rue Leo Saignat, F-33076 Bordeaux, France

Sek and Moore [J. Acoust. Soc. Am. **106**, 351–359 (1999)] and Lyzenga *et al.* [J. Acoust. Soc. Am. **116**, 491–501 (2004)] found that the just-noticeable frequency difference between two pure tones relatively close in time is smaller when these tones are smoothly connected by a frequency glide than when they are separated by a silent interval. This "glide effect" was interpreted as evidence that frequency glides can be detected by a specific auditory mechanism, not involved in the detection of discrete, time-delayed frequency changes. Lyzenga *et al.* argued in addition that the glide-detection mechanism provides little information on the direction of frequency changes near their detection threshold. The first experiment reported here confirms the existence of the glide effect, but also shows that it disappears when the glide is not connected smoothly to the neighboring steady tones. A second experiment demonstrates that the direction of a 750 ms frequency glide can be perceptually identified as soon as the glide is detectable. These results, and some other observations, lead to a new interpretation of the glide effect, and to the conclusion that continuous frequency changes may be detected in the same manner as discrete frequency changes.
© 2009 Acoustical Society of America.. [DOI: 10.1121/1.3050271]

I. INTRODUCTION

How do listeners detect continuous frequency changes? This is an important question because changes of that kind abound in speech and other meaningful sounds.

It is clear that, in some cases, continuous frequency changes may be detected by means of "static" spectral cues. Imposing a frequency modulation (FM) on a pure tone widens its power spectrum (Hartmann, 1997). This spectral-width cue will be usable, for instance, if the task is to detect a sinusoidal FM with a rate of several tens of hertz. A similar cue may be used if the task is to detect unidirectional frequency glides in tone bursts with a very short duration. However, since the spectral analysis of sounds by the auditory system is performed using filters with short impulse responses (Moore, 2004, Chaps. 1 and 5), the detection of *slow* FM in stimuli lasting several hundreds of milliseconds is presumably not based on static spectral cues. How is FM detected in such conditions?

The simplest hypothesis, often called "the snapshot hypothesis," is that FM is detected by means of comparisons between frequency samples taken at different times, as if the stimulus actually consisted of successive tone bursts. Hartmann and Klein (1980) proposed a mathematical model of

FM detection based on this assumption. They showed that the model correctly predicted, among other things, differences between the psychometric functions obtained for the detection of sinusoidal FM and for the discrimination between two successive steady tone bursts differing in frequency. In the same vein, Demany and Semal (1989) showed that the frequency dependence of thresholds in the sinusoidal-FM detection task can be accounted for on the basis of the snapshot hypothesis, at least up to 4 kHz.

An alternative hypothesis, on which we focus here, is that continuous frequency changes can be detected by a specific mechanism, not involved in the detection of frequency differences between temporally separate steady tones. This "dynamic mechanism" (in the words of Dooley and Moore, 1988) would encode FM as a primary feature of sounds. Such a view is consistent with the fact that, in the auditory cortex of mammals, many neurons respond in a strong and selective manner to frequency glides (e.g., Whitfield and Evans, 1965; Zhang *et al.* 2003).

About three decades ago, psychophysical support for the dynamic-mechanism hypothesis was looked for in experiments that aimed at demonstrating selective adaptation effects in the FM domain. The idea was that the dynamic-mechanism hypothesis would be supported if it appeared that the detection of a given FM was impaired by repeated previous presentations of the same FM with a larger modulation depth, while being less affected by previous presentations of

---

a)Author to whom correspondence should be addressed. Electronic mail: laurent.demany@u-bordeaux2.fr
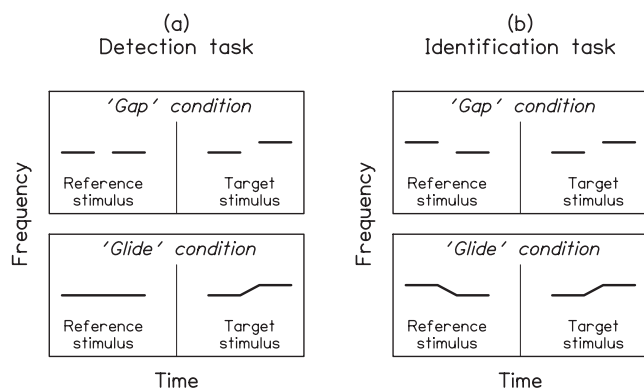
FIG. 1. Schematic spectrogram of stimuli used by Sek and Moore (1999) and Lyzenga *et al.* (2004). In the detection task, listeners had to discriminate target stimuli including a frequency change from reference stimuli with a steady frequency. In the identification task [not used by Sek and Moore (1999)], listeners had to discriminate target stimuli including a frequency rise from reference stimuli including a frequency fall. Both tasks were performed in a gap condition, where each stimulus consisted of two separate tone bursts, and in a glide condition where each stimulus had a continuous waveform.

stimuli including another type of FM, or amplitude modulation (AM) instead of FM. Several research groups did report selective adaptation effects of this type (for a review, see Kay, 1982). However, the data are now considered unconvincing, in part due to methodological problems (Wakefield and Viemeister, 1984), and also because the reported effects seem to disappear in trained listeners (Moody *et al.* 1984).

More recently, a quite different argument has been put forward in support of the dynamic-mechanism hypothesis. In two studies (Sek and Moore, 1999; Lyzenga *et al.* 2004), thresholds for the detection of frequency changes were measured using stimuli schematized in Fig. 1(a). On each trial, the listener was presented with two successive stimuli. One of them, the target stimulus, included a frequency change; the other (reference) stimulus did not. The task was to indicate if the target was the first or the second stimulus. In one condition, termed the "gap" condition, each stimulus consisted of two successive pure tones, with a silent gap in between. In a second condition, termed the "glide" condition, these two pure tones were no longer separated by a gap but smoothly connected to each other; so, the reference stimulus became a single pure tone and the target stimulus included a frequency glide smoothly connecting two frequency plateaux. For various durations of the gap or glide (5–200 ms) and of the frequency plateaux, it was found that the just-detectable frequency change was smaller in the glide condition than in the gap condition. To account for this finding, the authors argued that, in the glide condition, change detection was based on a combination of two cues: (1) a cue derived from a comparison between frequency samples taken at the beginning and the end of the stimulus, as in the gap condition; and (2) a cue provided by the glide itself. It could be reasonably assumed that the latter cue was not a static spectral cue. The authors thus interpreted this cue as the output of a change-detection mechanism responding exclusively to continuous or instantaneous changes.

In the two conditions described above, the listeners' task was merely to *detect* frequency changes. However, Lyzenga *et al.* (2004) also used the gap and glide conditions depicted in Fig. 1(b). On each trial, this time, both of the presented stimuli included a frequency change; the two changes had the same magnitude but opposite directions, and the task was to indicate if the upward change took place in the first or the second stimulus. Therefore, the listeners now had to *identify* the direction of frequency changes, which again varied in magnitude across trials. It appeared that, contrary to the detection thresholds, the identification thresholds were not significantly different in the gap and glide conditions. This led Lyzenga *et al.* (2004) to suggest that the dynamic change-detection mechanism provides little or no information about the direction of a frequency glide.

Here, we report new experiments that are closely related to those of Sek and Moore (1999) and Lyzenga *et al.* (2004). Experiment 1 tested the idea that the advantage of the glide condition over the gap condition in the detection task does not stem from the existence of a mechanism detecting the glide, but stems instead from the sole fact that the glide smoothly connects the two frequency plateaux. It was thought that this smooth connection could facilitate the detection of a difference between the plateaux, because the memorization of the first plateau might be improved by the transformation of a succession of two "auditory objects" into a single auditory object. Experiment 2 assessed the ability of listeners to identify the direction of both continuous and discrete frequency changes at their respective detection thresholds. According to the conclusions of Lyzenga *et al.* (2004), it should be difficult to identify the direction of a just-detectable frequency glide. In contrast, the snapshot hypothesis predicted that direction identification should not be more difficult for a glide than for a discrete change taking place following a gap.

II. EXPERIMENT 1

A. Method

1. Task and conditions

On each trial, two successive stimuli separated by a 700 ms interval were presented to the listener. One of them (the target) included a downward frequency change whereas the other (reference) included no frequency change. The listener had to indicate if the target was the first or the second stimulus, these two possibilities being a priori equiprobable. In order to force the listeners to base their judgments on *within-stimulus* frequency changes (see in this respect Sek and Moore, 1999), the center frequency of the stimuli was roved within trials. For each stimulus, this center frequency was selected randomly between 400 and 2400 Hz, the probability distribution being rectangular on a log-frequency scale. Four conditions, illustrated in Fig. 2, were run.

In condition 1, each stimulus consisted of two successive 250 ms sinusoidal tones, which had steady frequencies (differing from each other in the target stimulus) and were separated by a 250 ms silent gap. The tones had a nominal sound pressure level (SPL) of 65 dB and random initial phases. They were gated on and off with 10 ms cosinusoidal amplitude ramps.

Condition

FREQ.
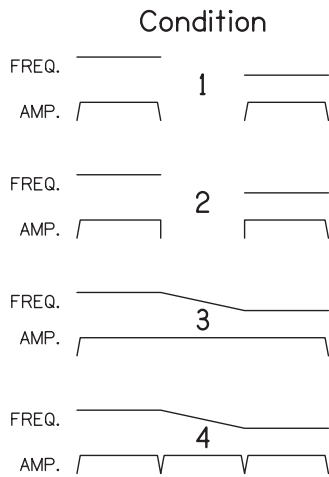AMP.
1

FREQ.
AMP.
2

FREQ.
AMP.
3

FREQ.
AMP.
4

FIG. 2. Schematic representation of the target stimuli used in the four conditions of experiment 1. In condition 1, each stimulus consisted of two successive sine tones, gated on and off with 10 ms ramps and separated by a silent gap; the second tone was lower in frequency than the first one. Condition 2 was identical except that the amplitude ramps limiting the central gap were much more abrupt. In condition 3, the central section of the stimuli was a falling frequency glide; the stimuli had a continuous waveform and their amplitude envelope was flat. In condition 4, frequency varied exactly as in condition 3, but the amplitude envelope was no longer flat because the three successive sections of the stimuli were gated on and off with 10 ms amplitude ramps. In all conditions, the amplitude ramps were cosinusoidal rather than linear.

Condition 2 was identical to condition 1 except for one point: In this second condition, the 250 ms gap located in the center of the stimuli was bounded by amplitude ramps of only 0.1 ms instead of 10 ms.

In condition 3, the gap was replaced, for the target stimulus, with a 250 ms frequency glide connecting 250 ms frequency plateaux without any change in amplitude or discontinuity in the stimulus waveform. This glide was linear on a logarithmic frequency scale. The reference stimulus was simply a 750 ms pure tone.

In the fourth and last condition, each stimulus consisted of three consecutive 250 ms tones, gated on and off with 10 ms cosinusoidal amplitude ramps. The onset ramps of the second and third tones started immediately after the end of the preceding offset ramps, so that the stimuli again had a total duration of 750 ms. The three tones making up the reference stimulus had identical steady frequencies. In the target stimulus, the first and third tones had steady but different frequencies, and the middle tone was a frequency glide starting with the frequency of the first tone and ending with the frequency of the third tone; as in condition 3, this glide was linear on a logarithmic frequency scale.

In each condition, change-detection thresholds were measured as described in Sec. II.A.2. Given the findings of Sek and Moore (1999) and Lyzenga *et al.* (2004), it was expected that thresholds would be lower in condition 3 than in conditions 1 and 2. With respect to condition 4, two opposite predictions could be made. If the advantage of condition 3 over conditions 1 and 2 stemmed from the detectability of a frequency change during the glides, then a reasonable prediction was that performance in condition 4 would be more similar to performance in condition 3 than to performance in conditions 1 and 2. Alternatively, it could be sup-

posed that the glides of condition 3 were advantageous not because they provided information but merely because they made the target stimuli continuous. Under this assumption, a logical prediction was that performance in condition 4 would be more similar to performance in condition 1 than to performance in condition 3. If any effect of discontinuity were due to the spectral splatter associated with the transition between sound and silence, rather than to the introduction of a silent gap, then performance in condition 2 might be somewhat worse than in condition 1 since the gap located in the center of the stimuli for these conditions had sharper bounds in condition 2 than in condition 1. The very sharp transitions of condition 2 produced more spectral splatter than the smoother transitions of condition 1.

2. Procedure and listeners

Thresholds were measured in separate blocks of trials for the four conditions, using an adaptive procedure tracking the 75% correct point on the psychometric function (Kaernbach, 1991). In each block of trials, the frequency change to be detected (C) initially had a magnitude of 60 cents (1 cent=1/100 semitone=1/1200 octave). C was decreased following each correct response, and increased following each incorrect response. A block ended after the 14th reversal in the variation of C. Up to the fourth reversal, C 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, C was either multiplied by 1.5 or divided by the cube root of this factor. The threshold measured in a block of trials was defined as the geometric mean of all the C values used from the fifth reversal onwards.

Listeners were tested individually in a triple-walled sound-attenuating booth (Gisol, Bordeaux). They wore headphones (Sennheiser HD265), through which the stimuli were delivered binaurally. The stimuli were generated via 24 bit digital-to-analog converters (RME) at a sampling rate of 44.1 kHz. Responses were given by means of mouse-clicks on two virtual buttons on a monitor screen, and were immediately followed by visual feedback. Response times were not limited. Within a block of trials, there was a pause of about 700 ms between each response and the onset of the next stimulus. Each experimental session consisted of two or three sequences of four blocks, within which each condition was used once, in a random position (1, 2, 3, or 4). Overall, the collected data consisted of 16 threshold measurements per condition and listener.

Five listeners were tested: four students who were in their twenties (L1, L2, L3, L4) and the first author (L5, 53). All listeners had normal audiograms up to 4 kHz and previous experience in similar tasks. They were given a few practice sessions before the experiment proper.

B. Results

Figure 3 displays the geometric means of the 16 threshold estimates made for each listener and condition (open symbols), as well as the grand geometric mean for each condition (filled circles). The geometric standard errors of the data points corresponding to the open symbols have an av-
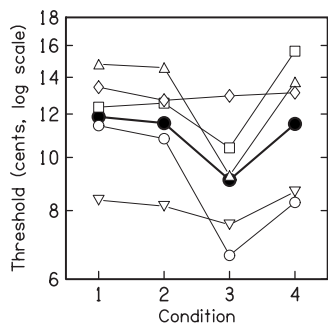
FIG. 3. Results of experiment 1. The open symbols represent the thresholds of the five listeners (L1: circles; L2: upward-pointing triangles; L3: downward-pointing triangles; L4: diamonds; L5: squares) for each condition. The filled circles represent the geometric means of these individual thresholds.

erage value of 8.8% (range: 4.3%–14.2%). It can be seen that, globally, performance was similar in conditions 1, 2, and 4, but better in condition 3. However, the effect of condition on performance was much larger for some listeners than for others. For each of the four planned pairwise comparisons (conditions 1 versus 2, 1 versus 3, 1 versus 4, and 3 versus 4), a two-way analysis of variance (ANOVA) (condition × listener) was performed on the logarithms of the threshold estimates. These ANOVAs confirmed that performance in condition 3 differed significantly from performance in condition 1 [$F(1,150)=26.1$, $P<0.001$] and condition 4 [$F(1,150)=24.5$, $P<0.001$], while performance in condition 1 did not differ from performance in conditions 2 and 4 ($F<1$ in each case). However, a significant interaction between the condition and listener factors was found in the comparisons between conditions 1 and 3 [$F(4,150)=3.9$, $P=0.005$], 1 and 4 [$F(4,150)=3.1$, $P=0.02$], and 3 and 4 [$F(4,150)=2.5$, $P=0.04$].

In order to check that the thresholds obtained in condition 4 were significantly more similar to those obtained in condition 1 than to those obtained in condition 3, we performed an additional ANOVA on the differences observed, within each sequence of four threshold measurements, between conditions 1 and 4, and between conditions 4 and 3; the processed data were again the logarithms of the threshold estimates. This two-way ANOVA [type of difference (1−4 versus 4−3) × listener] did reveal a significant main effect of type of difference [$F(1,150)=7.0$, $P=0.009$]; the interaction of the two factors was also found to be significant [$F(4,150)=2.9$, $P=0.025$].

C. Discussion

The fact that thresholds were significantly lower in condition 3 than in condition 1 confirms the main finding of Sek and Moore (1999) and Lyzenga et al. (2004). But on the other hand, their interpretation of this finding is seriously challenged by the fact that the mean threshold measured in our condition 4 was much more similar to the threshold obtained in condition 1 (or 2) than to the threshold obtained in condition 3. This observation indicates that the advantage of

condition 3 over condition 1 may stem merely from the temporal continuity of the stimuli used in condition 3 rather than from the existence of "glide detectors."

Actually, support for this hypothesis is provided by some of the results of Lyzenga et al. (2004). Their experiment on change detection included, in addition to the gap and glide conditions depicted in Fig. 1(a), a "noise" condition in which the central portion of the stimuli was neither a gap nor a frequency glide but a noise burst. This noise burst was presented at a relatively high level, and therefore elicited a continuity illusion. In the target stimuli, the listeners could hear illusory glides smoothly connecting the two frequency plateaux. Surprisingly, thresholds in this noise condition were significantly lower than thresholds in the gap condition. To account for that, the authors supposed that the dynamic mechanism detecting the real glides of the glide condition was also able to detect illusory glides. However, there is a somewhat circular aspect to this reasoning. When a frequency change was just-detectable in the noise condition, the same frequency change was not detectable in the gap condition. Hence, one would have to assume that the auditory system introduced an illusory glide between two tones whose frequencies it could not otherwise discriminate, and then used the glide to detect the frequency difference, rather like Baron Munchausen pulling himself out of a bog by his own hair. It seems more parsimonious to hypothesize that the noise condition was advantageous merely because in that condition the stimuli were perceived as continuous.

Returning to the results of the present experiment, one should note that the advantage of condition 3 over the other conditions might be ascribed to the use of spectral cues by the listeners. Although the target stimuli of condition 3 had a perfectly continuous waveform, some spectral splatter was produced in these stimuli when the initial frequency plateau suddenly became a frequency glide, and when the glide suddenly became a new plateau. This spectral splatter may have provided a cue since it was absent in the reference stimuli. In order to minimize the influence of that cue in their glide condition, Lyzenga et al. (2004) presented their stimuli in a background of pink noise. We did not do so. However, it will be seen that the results of our second experiment discredit the idea that the advantage of condition 3 stemmed from the use of spectral cues in this condition.

III. EXPERIMENT 2

A. Rationale

According to Lyzenga et al. (2004), the auditory system contains a mechanism specific for the detection of dynamic acoustic changes such as frequency glides, but this dynamic mechanism is not sensitive to the direction of a glide, or at least does not facilitate the identification of its direction. Such a view implies that listeners should be unable to identify the direction of a glide at its detection threshold. By contrast, as pointed out in the Introduction, a prediction of the snapshot hypothesis is that direction identification at the detection threshold should not be systematically more difficult for continuous changes than for discrete, time-delayed changes. In experiment 2, using both continuous frequency

changes and discrete ones, we compared the magnitude that a given type of change must have in order to be just-detectable to the magnitude that the same type of change must have in order to be reliably identified as an upward change or a downward change. One and the same psycho-physical paradigm was employed to measure the detection thresholds and the identification thresholds. This paradigm had been previously employed by Semal and Demany (2006) in a study investigating exclusively the perception of discrete changes.

B. Method

On each trial, as in experiment 1, the listener was pre-sented with two successive stimuli separated by a 700 ms interval: a target stimulus containing a frequency change and a reference stimulus in which frequency did not change. Again, the target stimulus was either the first or the second stimulus, equiprobably. This time, however, the direction of the frequency change was no longer fixed: Frequency could go up or down, equiprobably. In separate blocks of trials, which did not differ from each other with respect to the stimulus characteristics, the listeners performed two different tasks: a detection (D) task and an identification (I) task. In the D task, one had to indicate if the target was the first or the second stimulus. In the I task, one had to indicate if frequency changed upwards or downwards. As in experiment 1, the center frequency of the stimuli was roved within trials, and could take any value between 400 and 2400 Hz. In each block of trials, again, a fixed type of frequency change was produced, and the magnitude of the frequency change (in cents) was varied across trials in order to estimate a threshold defined as the magnitude of change for which the probability of a correct response was 0.75.

Five types of frequency change were used, yielding five experimental conditions. In condition 1, the stimuli were the same as those used in condition 1 of experiment 1, except for the randomization of change direction. In this condition, therefore, the D and I tasks were performed on stimuli in-cluding a central gap and the data provided information on the perception of discrete frequency changes. The frequency changes produced in the other four conditions were continu-ous. In condition 2, the stimuli were the same as those used in condition 3 of experiment 1, except for the randomization of change direction; each target stimulus thus consisted of a 250 ms frequency glide smoothly connecting 250 ms fre-quency plateaux, in a linear manner on a log-frequency scale. In the three remaining conditions, the target stimuli were "pure" frequency glides, in which frequency was constantly varying, linearly on a log-frequency scale. These glides, and the corresponding reference stimuli with a steady frequency, had a duration of 750 ms in condition 3, 250 ms in condition 4, and 50 ms in condition 5. In all conditions, the stimuli were gated on and off with 10 ms cosinusoidal amplitude ramps and had a nominal SPL of 65 dB.

The adaptive procedure used to measure thresholds was exactly the same as that employed in experiment 1, except for the initial magnitude of the frequency change presented to the listeners; this initial magnitude was 100 cents in the
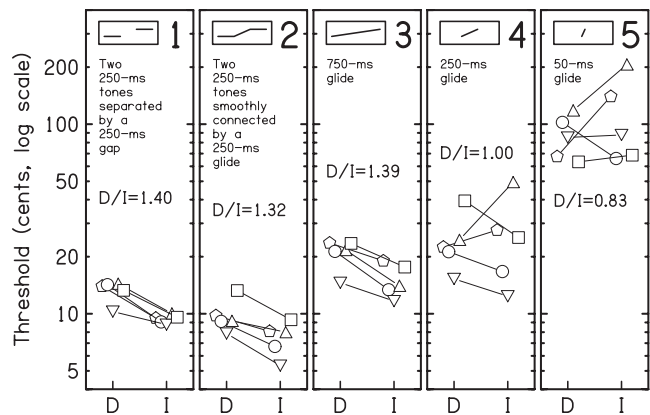

FIG. 4. Results of experiment 2. Detection (D) and identification (I) thresh-olds of the five listeners in the five conditions. For each condition, the quotient D/I of the geometric mean of the D thresholds and the geometric mean of the I thresholds is also indicated. The identical symbols represent the same listener in Figs. 3 and 4.

first four conditions, and 300 cents in condition 5 (because thresholds were markedly higher in the latter condition). In each experimental session, ten threshold measurements were made, one for each combination of condition and task (D or I). These ten measurements were made in a random order. Overall, the collected data consisted of 16 threshold mea-surements per condition, task, and listener.

We shall report here the data provided by five listeners. Four of them were identified as L1, L2, L3, and L5 in ex-periment 1; the fifth listener (L6), who did not participate in experiment 1, was an audiometrically normal student in her twenties. Four additional listeners were tested, but their re-sults will not be considered below because they had difficul-ties in the I task when the frequency changes were discrete (first condition).[1] Semal and Demany (2006) previously pointed out that, for some audiometrically normal listeners, it is difficult to identify the direction of very small but none-theless well-detected frequency changes between two succes-sive pure tones separated by a gap. Such listeners appeared to be inefficient *detectors* of frequency changes in Semal and Demany's study (2006).

C. Results and discussion

Figure 4 displays the geometric mean of the 16 threshold estimates made for each condition, task, and listener; in this figure and in Fig. 3, identical symbols represent the same listener. The geometric standard errors of the data points have an average value of 9.6% (range: 4.6%–18.3%). For each of the five conditions, we also indicate in Fig. 4 the statistic D/I obtained when the geometric mean of all the D thresholds is divided by the geometric mean of all the I thresholds.

In the first condition, involving discrete changes, each listener had a higher threshold in the D task than in the I task; D/I was equal to 1.40. An identical condition had been used by Semal and Demany (2006), and they obtained very similar results from their best three subjects (those who had the lowest thresholds in the two tasks). As explained in detail by Micheyl *et al.* (2008) (see also Semal and Demany, 2006), the standard version of signal detection theory—i.e., the

Demany *et al.*: Detection of continuous frequency changes

constant-variance Gaussian model—predicted in this condition (as well as the other four conditions used here) a D/I ratio of 1.56 for an ideal listener identifying always correctly the direction of a perceived frequency change. The empirically obtained D/I, 1.40, is close to this theoretical value. For an ideal listener identifying always correctly the direction of a perceived change, the "high-threshold" theory (Green and Swets, 1974, Chap. 5) predicted a D/I ratio of 1, i.e., a ratio *lower* than 1.40. Therefore, it is reasonable to conclude that the five listeners who provided the data analyzed here were able to identify the direction of a discrete frequency change as soon as this change was detectable.

Consider now the results obtained in condition 2, where the frequency changes were continuous. It can be noted first that each listener had a lower D threshold in this condition than in condition 1. This confirms the main finding of Sek and Moore (1999) and Lyzenga *et al.* (2004), as well as observations that we made in experiment 1. However, a more important result is that each listener also had a lower I threshold in condition 2 than in condition 1; the corresponding effect is statistically significant [$t(4) = 3.66$; $P = 0.02$]. From condition 1 to condition 2, D/I decreased, but only very slightly, as indicated in Fig. 4; this global trend was observed in only three of the five listeners. The fact that the I thresholds were lower in condition 2 than in condition 1 is not consistent with the idea that the advantage provided by the glides in condition 2 stems from the existence of a dynamic change-detection mechanism detecting glides without providing information on their direction. Another idea discredited by this finding is that, in the D task, the advantage of condition 2 over condition 1 could be merely due to the detection of spectral splatter in the transitions between the glides and the frequency plateaux flanking them: This spectral splatter provided no useful cue in the I task since it did not depend on the direction of the frequency changes (the ascending stimuli being temporal inversions of the descending stimuli).

In condition 3, the stimuli had the same duration as in conditions 1 and 2 (750 ms) and the targets had a constantly gliding frequency. As shown in Fig. 4, the D and I thresholds were consistently worse than in condition 1 (and 2), but D/I had almost exactly the same value as in condition 1. Both of these findings support the snapshot hypothesis. It was predicted by the snapshot hypothesis that thresholds would be worse than in condition 1 because taking a frequency sample of finite duration must result in some averaging of consecutive instantaneous frequencies; this averaging reduced the "effective" frequency span of the target stimuli in condition 3, but not condition 1. It is possible that some dynamic change-detection mechanism was sensitive to the frequency glides of condition 3 when they were above their detection threshold. However, our results suggest that at threshold, these glides were detected by means of snapshot comparisons.

In conditions 4 and 5, where the target stimuli consisted again of pure frequency glides but had shorter durations (250 and 50 ms, respectively), the thresholds increased, especially from condition 4 to condition 5. This trend is consistent with observations by Lyzenga *et al.* (2004). The I thresholds increased to a larger extent than the D thresholds, so that D/I decreased. This effect was strongly listener dependent, but it is clear that, overall, the shortening of stimulus duration reduced the listeners' ability to identify the direction of the frequency changes at their detection threshold. From the point of view of the snapshot hypothesis, the rise of the D and I thresholds makes sense: Detecting a frequency change within a short stimulus may be difficult because, in this case, the listener cannot sample frequency using a temporal window which is both short enough to allow for a comparison between nonoverlapping samples and long enough for accurate frequency measurements (Moore, 1973). In contrast, the reason why change detection should be more difficult for short stimulus durations is not obvious under the dynamic-mechanism hypothesis: Shortening a frequency glide spanning a fixed frequency distance increases, of course, the speed of frequency change; one might expect this increase in speed to have a favorable, rather than deleterious, effect for a dynamic change-detection mechanism.

We shall consider in Sec. IV how the decrease in D/I from condition 3 to condition 5 can be accounted for. The most important finding of the present experiment is that in conditions 2 and 3, as well as in condition 1, D/I was such that the listeners could be assumed to identify correctly the direction of a change as soon as they detected it. In conditions 2 and 3, our results are at odds with the idea that continuous (as opposed to discrete) frequency changes are optimally detected by an auditory mechanism providing no information about change direction. This idea had been put forth by Lyzenga *et al.* (2004) to account for their own data. In their study, however, D and I thresholds were not measured within the same experiment and compared to each other. Instead, Lyzenga *et al.* (2004) made only within-task (D or I) comparisons, as described in our Introduction and illustrated in Fig. 1. Consider again this figure. On each trial run in the I task, frequency changed by some amount a in the target stimulus, and $-a$ in the reference stimulus; this resulted in a difference of $2a$ between the two stimuli. In the D task, on the other hand, the corresponding difference was smaller by a factor of 2. So, for an ideal listener, at least in the gap condition, the threshold ratio D/I should have been equal to 2. An examination of the thresholds plotted in Figs. 2–5 of Lyzenga *et al.*, 2004 reveals that in the gap condition, D/I actually had a mean value of only 0.9. It thus seems that even in the gap condition, at least some of the listeners tested by Lyzenga *et al.* (2004) failed to identify correctly the direction of changes that they nonetheless reliably detected. Listeners showing this difficulty are not very uncommon (Semal and Demany, 2006). Indeed, as pointed out above, a sample of this population was tested in the present experiment. We did not report in detail the corresponding data here because they do not provide clear information concerning the main issue.

IV. EXPERIMENT 3

In experiment 2, D/I decreased from condition 3 to condition 5. One can make sense of this observation in the framework of the snapshot hypothesis by assuming that, in

TABLE I. Performance of the four subjects of experiment 3 in that experiment and in conditions 1 and 5 of experiment 2. These four listeners (L2, L3, L5, and L6) are, respectively, identified by upward-pointing triangles, downward-pointing triangles, squares, and pentagons in Fig. 4. The last column of the table displays the geometric means of the individual values.

	L2	L3	L5	L6	Mean
Experiment 3					
D threshold (cents)	93.1	60.6	71.1	67.8	72.2
D/I	0.35	1.19	1.41	0.65	0.79
Experiment 2, condition 5					
D threshold (cents)	118.2	85.1	63.3	67.5	81.0
D/I	0.57	0.97	0.92	0.48	0.70
Experiment 2, condition 1					
D threshold (cents)	14.4	10.3	13.3	14.0	12.9
D/I	1.43	1.17	1.39	1.47	1.36

the short target stimuli of conditions 4 and 5, the listeners sometimes misjudged the temporal order of frequency samples that they had correctly differentiated. We tested that idea in a small final experiment where the D and I tasks were performed on stimuli consisting of two successive 25 ms tones, gated on and off by means of 10 ms cosine ramps, with no silent interval in between. These two tones had steady frequencies, differing from each other in the target stimuli. The offset ramp of the first tone and the onset ramp of the second were sufficient to produce a well-audible gap between the two tones. The stimuli had the same overall duration (50 ms) as those used in condition 5 of experiment 2, but were structurally more similar to those of condition 1. Four listeners served as subjects. These were L2, L3, L5, and L6, the four listeners for whom, in condition 5 of experiment 2, D/I had been smaller than 1. L1 (identified by circles in Figs. 3 and 4) was not recruited again because in her case D/I had been similar in all conditions. Using the same procedure as before, 16 threshold measurements were made for each listener and task.

Table I displays the geometric means of these 16 measurements for the D task, and the associated values of D/I. We also indicate in this table the corresponding statistics for conditions 1 and 5 of experiment 2. It can be seen that for each listener, the D threshold measured in the present experiment was not very different from the D threshold measured previously in condition 5. With respect to D/I, however, there were substantial individual differences. For two listeners (L2 and L6), the new D/I was smaller than 1 and much closer to the value previously found in condition 5 than to the value found in condition 1. For the other two listeners (L3 and L5), the opposite was true.

The small D/I ratio observed here for L2 and L6 presumably originates from time-order confusions. For these two listeners, therefore, it may well be that the small D/I ratio obtained in condition 5 of experiment 2 also originated from time-order confusions. For L3 and especially L5, on the other hand, this is unlikely; their behavior here suggests that, in condition 5 of experiment 2, they did not detect the glides as specified by the snapshot hypothesis but used a different cue.[2] It is conceivable that the cue in question was the output of a dynamic change-detection mechanism which does not

provide optimal information about change direction. Such a mechanism would be more apt to detect frequency glides in short stimuli than in long ones because the slope of a just-detectable glide is larger when the stimulus is short, as shown by experiment 2 and several previous studies (Sergeant and Harris, 1962; Nabelek and Hirsh, 1969; Madden and Fire, 1997; Lyzenga et al., 2004). However, an alternative possibility is the existence of a static spectral cue. At any given time, the instantaneous excitation pattern produced by a gliding tone in a bank of auditory filters should be somewhat less sharp than if the tone had a steady frequency. This relative spreading of the "instantaneous internal spectrum" is presumably detectable per se when the slope of the glide exceeds a certain threshold.

V. GENERAL DISCUSSION

The present study confirmed the existence of an intriguing auditory phenomenon initially described by Sek and Moore (1999) and subsequently investigated by Lyzenga et al. (2004): A small frequency difference between two successive tones (long enough to evoke a maximally salient pitch) is easier to detect when the tones are smoothly connected by a frequency glide than when they are separated by a silent gap of the same duration. Our study also shows, however, that the origin of this glide effect may not be the one suggested up to now. It was previously thought that the glide improves the detection of a frequency change because a change can be heard during the glide itself. We suggest instead that the glide is advantageous just because it connects the two tones in a continuous manner, and thus transforms a succession of two auditory objects into a single auditory object. In support of this alternative hypothesis, we found in experiment 1 that the glide is no longer advantageous when its two ends are disconnected from the neighboring tones. Another observation supporting the same hypothesis was reported by Lyzenga et al. (2004): They found that when the glide is replaced with a noise burst producing a continuity illusion, change detection is still better than when the tones are separated by a silent gap.

The glide effect of Sek and Moore (1999) clearly deserves to be elucidated because, if its initial interpretation

Demany et al.: Detection of continuous frequency changes

were correct, this effect would be the strongest piece of psychophysical evidence for a dynamic change-detection mechanism in the auditory system. Our alternative interpretation implies that such a mechanism may, in fact, not exist. However, we did not demonstrate that continuous frequency changes are *always* encoded by means of snapshot comparisons or on the basis of static spectral cues. It is possible that the dynamic mechanism exists but detects a frequency glide only when the speed of frequency change exceeds some value, below which snapshot comparisons could be more efficient. Cusack and Carlyon (2003) showed that detecting a tone containing sinusoidal FM among several steady tones is easier than the reverse (detecting a steady tone in a background of FM tones). They took this as evidence that the auditory system encodes FM as a primary sound feature, or in other words that the dynamic mechanism exists. More recently, Carlyon *et al.* (2004) found that although listeners heard the FM associated with a sinusoidally modulated tone continue when a portion of that tone was replaced with noise, they could not tell whether or not the tone resumed at the same FM phase as that which it would have had if it had really remained on. Carlyon *et al.* (2004) suggested from this observation that the auditory system is able to encode sinusoidal FM in a way that discards information on the phase of the FM, while preserving information such as the carrier frequency and the presence, depth, and rate of FM. In both of these previous studies, the depth and rate of the FM used were such that the FM was well above its normal detection threshold. Correlatively, instantaneous frequency changed at a relatively high speed. Therefore, the conclusion drawn from the two studies is not inconsistent with the idea that no dynamic mechanism is involved in the detection of FM when the speed of frequency change is low. Alternative explanations are that explicit encoding of dynamic changes applies only to periodic FM, or that it is more important in tasks, like the ones employed by Cusack and Carlyon (2003), where the cognitive load is more demanding than the detection or discrimination of frequency changes applied to isolated tones.

Another possibility is that the dynamic mechanism does not detect efficiently isolated glides, such as those used here in conditions 3–5 of experiment 2, or condition 4 of experiment 1, but works better when the stimulus is a frequency glide immediately preceded and/or followed by a frequency plateau *without any discontinuity*. Lyzenga *et al.* (2004) attempted to account for the D thresholds obtained in their glide condition [Fig. 1(a), lower panel] by assuming that in this condition the listeners combined two cues: a cue derived from a comparison between the two frequency plateaux and a cue provided during the glide itself by a dynamic change-detection mechanism. The first cue was assessed from the D thresholds measured in the gap condition, and the second from the D thresholds measured in a condition where the target stimulus consisted of nothing more than a frequency glide.[3] Lyzenga *et al.* (2004) found that performance in their glide condition was too good to be predicted in a simple manner from performance in the other two conditions. They assumed, therefore, that the two change-detection mechanisms providing the cues allegedly combined in the glide condition operated in a synergistic way. This apparent synergy was taken as evidence that frequency glides are more efficiently detected when they are preceded and/or followed by a frequency plateau. However, using the same equations as Lyzenga *et al.* (2004) we failed to find evidence for a synergistic effect in the results of the second experiment reported here: The D thresholds in condition 2 could be predicted from the D thresholds in conditions 1 and 4 by assuming simply that the internal noise limiting performance in condition 2 had two partially independent sources, respectively limiting performance in condition 1 and in condition 4. Under our hypothesis about the origin of the glide effect, the success of this prediction is fortuitous and listeners were, in fact, not using in condition 2 the information used in condition 4. It is warranted to assume that listeners extracted no information from the glides in condition 2 since they clearly extracted no information from the glides in condition 4 of experiment 1.

Although the results reported here are compatible with the idea that the auditory system contains a dynamic change-detection mechanism detecting efficiently frequency glides *when they are smoothly connected to frequency plateaux*, we think that our alternative interpretation of Sek and Moore's (1999) glide effect is more parsimonious. Admittedly, this alternative interpretation based on the snapshot hypothesis is incomplete: It remains to be explained why detecting a small difference between two frequency samples should be easier when they belong to the same auditory object than when this is not the case. But the snapshot hypothesis itself is obviously very reasonable. An important point is that snapshot comparisons are not necessarily less "automatic" than the dynamic change-detection mechanism suggested by Sek and Moore (1999) and Lyzenga *et al.* (2004). Indeed, Demany and Ramos (2005, 2007) (see also Demany *et al.*, 2008) showed that a frequency difference between two pure tones separated by a substantial time interval can be consciously perceived as an upward or downward pitch change even when the first of these tones has not been consciously perceived. This phenomenon strongly suggests that the auditory system contains automatic frequency-shift detectors. It may well be that these detectors respond not only to discrete, time-delayed frequency shifts, but also to continuous frequency changes.

ACKNOWLEDGMENTS

[1]We knew in advance that, for these four listeners, in the first condition, the I threshold would be markedly higher than the D threshold. This group was nevertheless tested in order to see if, for some listeners, it can be easier to identify the direction of a just-detectable frequency change when the change is continuous than when it is discrete. We did not observe significant trends in that direction. The four listeners' performance was markedly poorer in the I task than in the D task for all conditions, not only the first one. In this group, the ratios of the D and I thresholds were strongly correlated across conditions. In all conditions, the D thresholds were also systematically higher than those measured in the five listeners who provided the data analyzed below.

[2]For L3 and L5, D/I was larger than 1 in condition 4 of experiment 2,

where the target stimuli consisted of 250 ms frequency glides. Thus, the "different cue" hypothetically used by L3 and L5 was presumably used only for the 50 ms glides of condition 5.

[3]Lyzenga *et al.* (2004) used the classical model of signal detection theory to compute their predictions. Sek and Moore (1999) had previously analyzed their own data in a similar way, but their predictions were problematic because some of the relevant data were actually missing.

Carlyon, R. P., Micheyl, C., Deeks, J. M., and Moore, B. C. J. (**2004**). "Auditory processing of real and illusory changes in FM phase," J. Acoust. Soc. Am. **116**, 3629–3639.

Cusack, R., and Carlyon, R. P. (**2003**). "Perceptual asymmetries in audition," J. Exp. Psychol. Hum. Percept. Perform. **26**, 713–725.

Demany, L., Pressnitzer, D., and Semal, C. (**2008**). "On the binding of successive tones: Implicit versus explicit pitch comparisons," J. Acoust. Soc. Am. **123**, 3049.

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 Ramos, C. (**2007**). "A paradoxical aspect of auditory change detection," in *Hearing—From Sensory Processing to Perception*, edited by B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey (Springer, Heidelberg), pp. 313–321.

Demany, L., and Semal, C. (**1989**). "Detection thresholds for sinusoidal frequency modulation," J. Acoust. Soc. Am. **85**, 1295–1301.

Dooley, G. J., and Moore, B. C. J. (**1988**). "Detection of linear frequency glides as a function of frequency and duration," J. Acoust. Soc. Am. **84**, 2045–2057.

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

Hartmann, W. M. (**1997**). *Signals, Sound, and Sensation* (AIP, Woodbury, New York).

Hartmann, W. M., and Klein, M. A. (**1980**). "Theory of frequency modulation detection for low modulation frequencies," J. Acoust. Soc. Am. **67**, 935–946.

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

Kay, R. H. (**1982**). "Hearing of modulation in sounds," Physiol. Rev. **62**, 894–975.

Lyzenga, J., Carlyon, R. P., and Moore, B. C. J. (**2004**). "The effects of real and illusory glides on pure-tone frequency discrimination," J. Acoust. Soc. Am. **116**, 491–501.

Madden, J. P., and Fire, K. M. (**1997**). "Detection and discrimination of frequency glides as a function of direction, duration, frequency span, and center frequency," J. Acoust. Soc. Am. **102**, 2920–2924.

Micheyl, C., Kaernbach, C., and Demany, L. (**2008**). "An evaluation of psychophysical models of auditory change perception," Psychol. Rev., **115**, 1069–1083.

Moody, D. B., Cole, D., Davidson, L. M., and Stebbins, W. C. (**1984**). "Evidence for a reappraisal of the psychophysical selective adaptation paradigm," J. Acoust. Soc. Am. **76**, 1076–1079.

Moore, B. C. J. (**1973**). "Frequency difference limens for short-duration tones," J. Acoust. Soc. Am. **54**, 610–619.

Moore, B. C. J. (**2004**). *An Introduction to the Psychology of Hearing*, 5th ed. (Elsevier, Amsterdam).

Nabelek, I. V., and Hirsh, I. J. (**1969**). "On the discrimination of frequency transitions," J. Acoust. Soc. Am. **45**, 1510–1519.

Sek, A., and Moore, B. C. J. (**1999**). "Discrimination of frequency steps linked by glides of various durations," J. Acoust. Soc. Am. **106**, 351–359.

Semal, C., and Demany, L. (**2006**). "Individual differences in the sensitivity to pitch direction," J. Acoust. Soc. Am. **120**, 3907–3915.

Sergeant, R. L., and Harris, J. D. (**1962**). "Sensitivity to unidirectional frequency modulation," J. Acoust. Soc. Am. **34**, 1625–1628.

Wakefield, G. H., and Viemeister, N. F. (**1984**). "Selective adaptation to linear frequency-modulated sweeps: Evidence for direction-specific FM channels?," J. Acoust. Soc. Am. **75**, 1588–1592.

Whitfield, I. C., and Evans, E. F. (**1965**). "Responses of auditory cortical neurons to stimuli of changing frequency," J. Neurophysiol. **28**, 655–672.

Zhang, L. I., Tan, A. Y. Y., Schreiner, C. E., and Merzenich, M. M. (**2003**). "Topography and synaptic shaping of direction selectivity in primary auditory cortex," Nature (London) **424**, 201–205.