

# The role of peripheral resolvability in pitch-sequence processing

**Marion Cousineau<sup>a)</sup>**

*Laboratoire Psychologie de la Perception, CNRS and Université Paris Descartes, 45 rue des Saints-Pères, F-75006 Paris, France and Département d'études cognitives, Ecole normale supérieure, 29 rue d'Ulm, F-75005 Paris, France  
marion.cousineau@ens.fr*

**Laurent Demany**

*Laboratoire Mouvement, Adaptation, Cognition (UMR CNRS 5227), Université de Bordeaux, 146 rue Leo Saignat, BP 63, F-33076 Bordeaux, France  
laurent.demany@psyac.u-bordeaux2.fr*

**Daniel Pressnitzer**

*Laboratoire Psychologie de la Perception, CNRS and Université Paris Descartes, 45 rue des Saints-Pères, F-75006 Paris, France and Département d'études cognitives, Ecole normale supérieure, 29 rue d'Ulm, F-75005 Paris, France  
daniel.pressnitzer@ens.fr*

**Abstract:** The authors previously reported that same/different judgments on pitch sequences were more accurate for tones with resolved (low-rank) harmonics compared to unresolved (high-rank) harmonics, even when discriminability between tones was equated [Cousineau *et al.* (2009). *J. Acoust. Soc. Am.* **126**, 3179–3187]. Here, peripheral resolvability, defined by the number of harmonics per cochlear filter, was contrasted with harmonic number. Tones were presented either diotically or dichotically. In the latter case, even and odd harmonics were presented to different ears, thus halving the number of harmonics per cochlear filter. Performance was better for dichotic than for diotic presentations. This indicates that peripheral resolvability is necessary and sufficient for efficient pitch-sequence processing.

© 2010 Acoustical Society of America

**PACS numbers:** 43.66.Mk, 43.66.Hg [QJF]

**Date Received:** August 3, 2010    **Date Accepted:** September 16, 2010

## 1. Introduction

Complex periodic sounds, such as notes from musical instruments or vowels in speech, can be decomposed into a series of harmonics that are integer multiples of a common fundamental frequency ( $F_0$ ). When all harmonics are present, the rank of the harmonic (2nd for  $2F_0$ , 3rd for  $3F_0$ , etc.) is indicative of the number of harmonics within a peripheral cochlear filter. This is because the spacing between successive harmonics is constant and equal to  $F_0$ , whereas the frequency selectivity of the cochlea is approximately logarithmic. When a harmonic is processed independently within a filter, it is said to be peripherally “resolved.” Conversely, when several harmonics interact within a filter, they are “unresolved” (Shackleton and Carlyon, 1994). Even though the precise criterion for resolvability remains a matter of controversy, it has been repeatedly observed that  $F_0$  difference limens increase sharply around the 10th harmonic (e.g., Houtsma and Smurzynski, 1990; Shackleton and Carlyon, 1994; Krumbholz *et al.*, 2000). Recently, we have shown that same/different judgments on sequences of tones varying in  $F_0$  were also more accurate for peripherally resolved, low-rank harmonics than for peripherally unresolved, high-rank harmonics (Cousineau *et al.*, 2009). Importantly, the method used in

---

<sup>a)</sup> Author to whom correspondence should be addressed.

these experiments controlled for the discriminability between tones, so that the observed “sequence processing” advantage for resolved harmonics cannot be attributed to their better F0 difference limens.

It is not clear yet whether the sequence processing advantage depends on peripheral resolvability or harmonic number. Bernstein and Oxenham (2003) introduced a method to distinguish between the two. They used harmonic complexes with even and odd harmonics presented to opposite ears, effectively halving the number of harmonics per cochlear filter compared to diotic presentation. It should be noted here that pitch perception across the ears is possible. Indeed, a mistuned harmonic of a complex to which it is presented contralaterally has a mandatory effect on the overall perceived pitch (e.g., Gockel *et al.*, 2005).

Bernstein and Oxenham (2003) asked listeners to make frequency comparisons between a pulsed harmonic within a complex and a pure tone probe presented before the complex. Dichotic presentation improved performance, consistent with the improved peripheral resolution due to dichotic presentation. Bernstein and Oxenham (2003) then compared F0 difference limens for diotic and dichotic harmonic complexes. Surprisingly, dichotic presentation did not improve F0 difference limens. The authors concluded that the mechanism of residue pitch extraction was not able to benefit from peripheral resolvability of high-rank harmonics. Experimental details seem important to observe the effect. Using a slightly different technique (fixed ear of presentation and harmonic numbers within a trial, see their Footnote 1 and our Discussion), Bernstein and Oxenham (2008) did observe a benefit of dichotic presentation for F0 difference limens.

Here, we tested sequence processing (Cousineau *et al.*, 2009) using stimuli inspired by Bernstein and Oxenham (2003, 2008). A diotic condition used unresolved harmonics (above the 10th), for which poor sequence processing is expected (Cousineau *et al.*, 2009). A dichotic condition presented even and odd harmonics to different ears, in order to re-establish peripheral resolvability while leaving the harmonic ranks unchanged. A third condition used monaural presentation of the even harmonics only. If peripheral resolvability is sufficient to support efficient sequence processing, the dichotic and monaural conditions should display better performance than the diotic condition. Alternatively, if harmonic rank is the relevant parameter, then only the monaural condition should display good performance.

## 2. Methods

### 2.1 Stimuli

The sequence elements were bandpass-filtered harmonic complex tones with a F0 close to 125 Hz. The bandpass filter was held constant throughout the experiment. The edges of the filter were rounded with a quarter cycle of a cosine function (Krumbholz *et al.*, 2000) to avoid the perception of edge tones (Kohlrausch *et al.*, 1992). The cutoffs of the frequency band (−3 dB points) were 1375 and 2375 Hz (Shackleton and Carlyon, 1994), so that harmonics 11–19 of a 125-Hz complex would fall within the flat portion of the filter. The lower and the upper spectral ramps had widths of 0.2 and 1.0 kHz, respectively. The relative phase of the components was random. In order to mask distortion products (Pressnitzer and Patterson, 2001), the tones were mixed with pink noise (from 62.5 to 22050 Hz) in all conditions. For each stimulus, the overall level of the noise was set at 6 dB below the overall level of the tone; then, the sound pressure level (SPL) of the tone plus noise compound was set close to 65 dB. This yielded a signal-to-noise ratio of 9.0 dB for each harmonic relative to its surrounding third-octave band of noise. Stimuli were 200-ms long (including 25-ms on and off raised-cosine ramps).

The harmonic complexes could be presented (1) diotically, with all harmonics presented to both ears (condition “Dio”), (2) dichotically, with even and odd harmonics presented to opposite ears (condition “Dicho”), or (3) monaurally, with only even harmonics presented to one ear (condition “Mono”). In the Dicho and Mono conditions, the frequency spacing of the harmonics is doubled in each ear. The masking noise was identical in the two ears for the Dio and Dicho conditions; it was monaural in the Mono condition.

## 2.2 Procedure

Six listeners with no self-reported hearing disorder participated in the experiment (mean age = 26.0, SD = 3.2, 3 female). Half of them received the even harmonics of the dichotic condition and the monaural stimulation in the right ear throughout the experiment, and the other half in the left ear.

The procedure has been described in detail elsewhere (Cousineau *et al.*, 2009, 2010). Briefly, it comprised a preliminary adjustment phase designed to factor out tone discriminability, and a main testing phase. In the preliminary adjustment phase, listeners had to perform a same/different task on two tones with variable differences in F0. The reference tone had an F0 of 125 Hz, while the other tone had an F0 of  $125 + \Delta$  Hz. For each condition (Dio, Dicho, or Mono) and listener, we estimated the stimulus difference  $\Delta$  yielding a  $d'$  value of 2.

In the main testing phase, the previously determined values of  $\Delta$  were used in binary sequences of  $N = 1, 2,$  or  $4$  tones. On each trial, two sequences separated by a 400-ms silence were presented. In the first sequence, each tone was, at random, either a reference stimulus, A, with a 125-Hz F0, or another stimulus, B, differing positively from A by  $\Delta$ . The second sequence was equiprobably identical to the first sequence or different from it. In the latter case, a single, randomly chosen tone was changed from A to B or vice versa. The listener had to make a same/different judgment. For each listener, condition, and  $N$  value, four blocks of 50 trials were run. The ordering of conditions and  $N$  values was randomized. No training was provided apart from the series of adjustment blocks, which corresponds to sequences with  $N = 1$ .

Stimuli were generated on a personal computer and then played through an RME Fireface audio sound card and digital to analog converter, with 16-bit coding accuracy and 44.1 kHz sampling rate. They were delivered by means of Sennheiser HD 250 Linear II headphones. Listeners were seated in a double-walled sound insulated booth (IAC). Responses were given by means of button presses. No feedback was provided.

## 3. Results

### 3.1 Adjustment step

$\Delta$  was smallest in the Mono condition, with a mean value of 1.43% of F0 and a standard error (SE) across subjects of 0.2%. The mean  $\Delta$  values for conditions Dicho and Dio were respectively 2% (SE: 0.2%) and 20.5% (SE: 5.7%). A repeated-measures ANOVA on the log-transformed  $\Delta$  values revealed a significant main effect of Condition [ $F(2, 10) = 141.6, P < 0.0001$ ]. Bonferroni-corrected paired  $t$ -tests were used to further evaluate the effect of Condition on  $\Delta$  values (all displayed  $P$  values for  $t$ -tests are corrected). They revealed significant differences between the  $\Delta$  values for Mono and Dio ( $P = 0.0002$ ), for Mono and Dicho ( $P = 0.005$ ), and for Dio and Dicho ( $P = 0.0002$ ).

### 3.2 Sequence discrimination task

The left panel of Fig. 1 shows the effect of  $N$  and condition on  $d'$ . For  $N = 1$ ,  $d'$  was relatively similar in all three conditions. Slightly better  $d'$  values on average were obtained for the Dio condition, indicating that the adjustment phase had not been completely accurate. The pattern of results as a function of  $N$  differed markedly between conditions. For condition Dio, performance steadily decreased with  $N$ . For conditions Dicho and Mono, performance did not decrease from  $N = 1$  to  $N = 2$  and decreased only modestly from  $N = 2$  to  $N = 4$ .

A repeated-measures ANOVA ( $N \times$  Condition) revealed a significant main effect of  $N$  [ $F(2, 10) = 10.75, P = 0.003$ ]. The main effect of Condition was not significant [ $F(2, 10) = 0.08, P = 0.92$ ]. However, a significant interaction between the two experimental factors was found [ $F(4, 20) = 3.03, P = 0.04$ ]. This confirms that the effect of  $N$  was not the same in all conditions.

We performed an additional analysis to summarize the data. Straight lines were fitted to the individual data obtained for  $N = 1, 2$  and  $4$ , using a log-scale for  $N$ . The slope of the fitted lines characterizes the effect of  $N$  on performance, irrespective of performance for  $N = 1$ . We call this statistic the “ $d'$  slope” (Cousineau *et al.*, 2009, 2010). The right panel of Fig. 1 displays

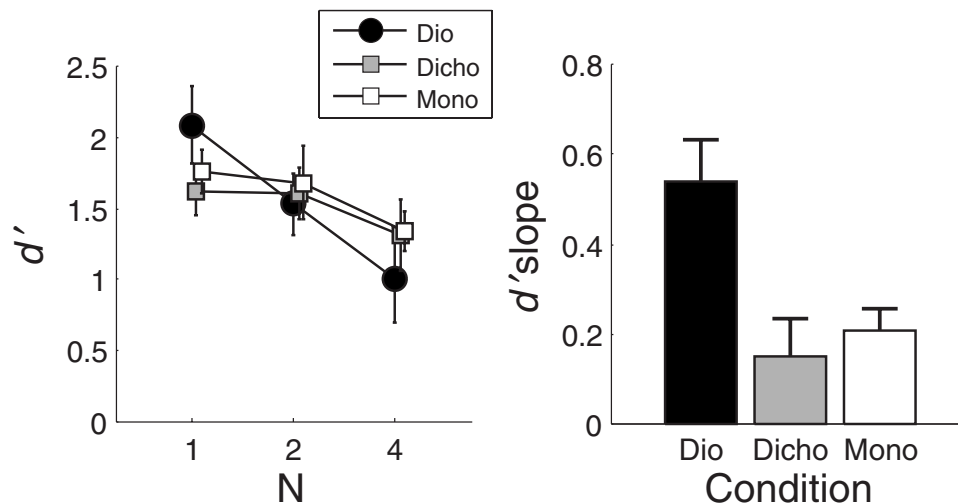


Fig. 1. Mean results of the main experiment. Left: Mean value of  $d'$  as a function of  $N$  in the three experimental conditions. Error bars are standard errors of the means. Right: Mean and standard error of the " $d'$  slope," summarizing the decrease in performance from  $N=1$  to  $N=4$ .

the mean  $d'$  slopes for the three conditions. The mean slope found for Dio was the highest, indicating poorer sequence processing. Mean slopes for Dicho and Mono were markedly smaller.

A repeated-measures ANOVA of the  $d'$  slopes revealed a significant main effect of Condition [ $F(2, 10) = 12.0, P < 0.002$ ]. Bonferroni-corrected paired  $t$ -tests showed that the  $d'$  slope for Dio differed significantly from those for Dicho and Mono ( $P < 0.05$  in both cases), but that  $d'$  slopes for Dicho and Mono did not differ significantly from each other ( $P = 0.90$ ). Figure 2 displays individual results for the  $d'$  slope. As can be seen in the left panel, all but one listener displayed a steeper slope for Dio than for Mono. In contrast, when Dicho and Mono are compared (right panel), all but one listener sit on the diagonal. The latter finding demonstrates a remarkably similar sequence processing ability for these two conditions when individual data are considered.

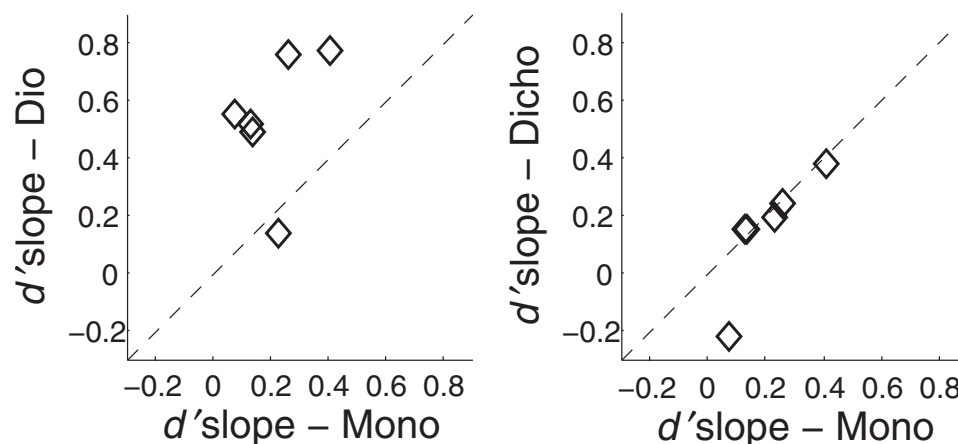


Fig. 2. Individual results obtained in the main experiment. Left: Individual data for the  $d'$  slope (see text) in the Mono and Dio conditions. Each diamond represents one listener. Right: Same, but for the Mono and Dicho conditions.

#### 4. Supplementary experiment

Since the adjustment phase aimed at equating discriminability between two tones in the different conditions, the resulting  $\Delta$  values can be considered as F0 difference limens. In sharp contrast with [Bernstein and Oxenham \(2003\)](#), but in agreement with a later report by [Bernstein and Oxenham \(2008\)](#), we observed a large improvement for  $\Delta$  in dichotic presentation compared to diotic presentation. However, our adjustment was not perfect, and there were methodological differences between our study and those of [Bernstein and Oxenham \(2003, 2008\)](#). In this supplementary experiment, F0 difference limens were measured for the three conditions of the main experiment, in the same six subjects but with a procedure similar to [Bernstein and Oxenham \(2008\)](#): a three-alternative forced-choice procedure with a 2-down/1-up stepping rule, ear of presentation and harmonic numbers being fixed within trials. Five threshold estimations were obtained for each subject and condition; the mean of the last four was taken as the final estimation.

Performance was best in the Mono condition (mean: 0.8%, SE: 0.1%) and worst in the Dio condition (mean: 17.7%, SE: 8.1%). Intermediate results were obtained for Dicho (mean: 1.4%, SE: 0.2%). An ANOVA on the log-transformed data revealed a significant main effect of condition [ $F(2, 10)=28.5$ ,  $P < 0.0001$ ]. Bonferroni-corrected paired  $t$ -tests revealed significant differences between Dio and Dicho conditions ( $P=0.004$ ), Dio and Mono conditions ( $P=0.002$ ), and between Dicho and Mono conditions ( $P=0.02$ ). This confirms that with our stimuli and task, dichotic presentation of the odd and even harmonics improves F0 difference limens.

#### 5. Discussion

This study demonstrated that efficient processing of pitch sequences is achieved when the harmonics forming the tones are peripherally resolved, even for harmonics above the 10th. Peripheral resolvability was obtained by dichotic presentation of the harmonics, with even and odd harmonics in different ears. This dichotic presentation markedly improved sequence processing, compared to a diotic presentation. In our Dicho condition, nonetheless, listeners could not totally ignore the ear receiving the odd harmonics, as shown by the fact that F0 difference limens were slightly but significantly poorer than in the Mono condition.

Previously, we argued that frequency-shift detectors (FSDs, [Demany and Ramos, 2005](#)) were the likely mechanism for efficient pitch-sequence processing ([Cousineau \*et al.\*, 2009](#)). If this is the case, one might infer from the present findings that the FSDs operate before the binaural combination of information from the two ears, and can thus benefit from peripheral resolvability before binaural convergence. Alternatively, it may be that the FSDs operate after binaural convergence but that precise frequency information is not lost in the combination process. [Demany and Ramos \(2007\)](#) reported data that are consistent with the latter hypothesis. They tested the capacity of listeners to identify the direction of a frequency change between: (i) a pure tone informationally masked by a simultaneous chord; (ii) a probe tone following the chord. Performance was barely poorer when (i) and (ii) were presented to contralateral ears than when they were presented to the same ear. This shows that the FSDs must be located after binaural convergence.

At first sight, our results conflict with the findings of [Bernstein and Oxenham \(2003\)](#), who suggested that peripheral resolvability alone is not sufficient to guarantee accurate F0 discrimination. On the other hand, our results are consistent with an observation reported later by [Bernstein and Oxenham \(2008, Experiment 1\)](#). It is worth considering the similarities and differences between the three studies. All studies used similar harmonic complex tones and procedures (cf. our Supplementary experiment). However, we used masking noise correlated between the ears whereas [Bernstein and Oxenham \(2003\)](#) used uncorrelated noise. Uncorrelated noise might induce a binaural signal-detection advantage in the diotic condition that could explain part of the discrepancy. Perhaps more importantly, the stimuli differed in the way the contribution of edge tones was minimized. Whereas we used rounded spectral edges, [Bernstein and Oxenham \(2003\)](#) varied the lowest harmonic within each trial (for a nominal harmonic number  $n$ , the 3 intervals of a trial randomly contained harmonics  $n-1$  to  $n+10$ ,  $n$  to  $n+11$ , and

$n+1$  to  $n+12$ ). With the latter method, timbre differences between the stimuli are introduced, and the frequency shifts of the lowest and highest harmonics are often not congruent with the F0 changes. Thus, listeners have to ignore timbre differences or frequency shifts of individual harmonics, and are forced to base their responses on F0 per se (Michey *et al.*, in press). In contrast, our listeners as well as those of Bernstein and Oxenham (2008, Experiment 1) could have performed the F0 discrimination task on the basis of frequency shifts of individual harmonics, without necessarily extracting an accurate estimate of overall F0. If this were the case, a parsimonious account of all three studies would be that F0 pitch estimation depends on harmonic rank, but that FSDs can operate on a representation of individual frequencies that is accurate even after binaural convergence, as long as peripheral resolvability is ensured. The latter representation could be similar to the central spectrum proposed by Srulovicz and Goldstein (1983).

## 6. Conclusion

The present data show that accurate same/different judgments for pitch sequences are observed for high-rank harmonics that are peripherally resolved because of dichotic presentation. This finding, together with the data of Cousineau *et al.* (2009, 2010), show that peripheral resolvability is necessary and sufficient for accurate pitch sequence processing. Additionally, we suggest that the encoding of frequency shifts and the estimation of F0 pitch recruit relatively independent mechanisms, which are likely to be combined for efficient pitch sequence processing in natural listening situations. For instance, these two mechanisms could respectively form the basis of contour and interval perception in melodies. Further work is needed to test this hypothesis.

## Acknowledgment

This work was partly supported by a grant from the FRM (Fondation pour la Recherche Médicale).

## References and links

- Bernstein, J. G., and Oxenham, A. J. (2003). "Pitch discrimination of diotic and dichotic tone complexes: harmonic resolvability or harmonic number?," *J. Acoust. Soc. Am.* **113**, 3323–3334.
- Bernstein, J. G., and Oxenham, A. J. (2008). "Harmonic segregation through mistuning can improve fundamental frequency discrimination," *J. Acoust. Soc. Am.* **124**, 1653–1667.
- Cousineau, M., Demany, L., Meyer, B., and Pressnitzer, D. (2010). "What breaks a melody: Perceiving pitch and loudness sequences with a cochlear implant," *Hear. Res.* **269**, 34–41.
- Cousineau, M., Demany, L., and Pressnitzer, D. (2009). "What makes a melody: The perceptual singularity of pitch sequences," *J. Acoust. Soc. Am.* **126**, 3179–3187.
- 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.
- Gockel, H., Carlyon, R. P., and Plack, C. J. (2005). "Dominance region for pitch: Effects of duration and dichotic presentation," *J. Acoust. Soc. Am.* **117**, 1326–1336.
- Houtsma, A. J. M., and Smurzynski, J. (1990). "Pitch identification and discrimination for complex tones with many harmonics," *J. Acoust. Soc. Am.* **87**, 304–310.
- Kohlrausch, A., Houtsma, A. J., and Evans E. F. (1992). "Pitch related to spectral edges of broadband signals [and discussion]," *Philos. Trans. R. Soc. London, Ser. B* **336**, 375–382.
- Krumbholz, K., Patterson, R. D., and Pressnitzer, D. (2000). "The lower limit of pitch as determined by rate discrimination," *J. Acoust. Soc. Am.* **108**, 1170–1180.
- Michey, C., Divis, K., Wroblewski, D. M., and Oxenham, A. J. "Does fundamental-frequency discrimination measure virtual pitch discrimination?," *J. Acoust. Soc. Am.* (in press).
- Pressnitzer, D., and Patterson, R. D. (2001). "Distortion products and the pitch of harmonic complex tones," in *Physiological and Psychophysical Bases of Auditory Function*, edited by A. J. M. Houtsma, A. Kohlrausch, V. F. Prijs, and R. Schoonhoven (Shaker, Maastricht), pp. 97–104.
- Shackleton, T. M., and Carlyon, R. P. (1994). "The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination," *J. Acoust. Soc. Am.* **95**, 3529–3540.
- Srulovicz, P., and Goldstein, J. L. (1983). "A central spectrum model: A synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum," *J. Acoust. Soc. Am.* **73**, 1266–1276.