A model of the pitch shifts of mistuned partials

Alain de Cheveigné (CNRS/Université Paris 7 and ATR HIP), and Minoru Tsuzaki (ATR HIP)

1 Abstract

Pitch shifts of partials mistuned from a harmonic series are explained by supposing that an internal randomly distributed variable determines both the pitch and the probability of fusion within the harmonic complex. When the variable falls near the harmonic series, the partial tends to fuse within the complex and its pitch is difficult to match. The distribution of successful pitch matches is therefore distorted in the direction of mistuning, hence the pitch shift. The model accounts for the major aspects of shifts observed experimentally.

2 Introduction

Hartmann et al. (1990) found that partials of a harmonic complex are easier to hear out when they are mistuned. They also noticed that subjects tended to overestimate the partial's mistuning when matching its pitch to a pure tone. This was confirmed by Hartmann and Doty (1996). The pitch shift had the same sign as the mistuning, and usually peaked at 4% mistuning and decreased beyond.

On the basis of these shifts, Hartmann and Doty (1996) argued against the model of Terhardt (1979) that predicts shifts in one direction whatever the mistuning. They proposed instead a time-domain model based on peaks of interspike interval (ISI) histograms. The model successfully accounted for major aspects of the shifts (direction, magnitude) of all components of rank greater than 1. However it could not account for shifts observed at the fundamental, nor could it account for the saturation and decrease of the pitch shift beyond 4% mistuning.

Here we examine a different explanation, based on harmonic fusion. Partials that match the harmonic series of a complex tone tend to fuse with it, whereas mistuned partials segregate and are easier to hear (Moore et al. 1985, 1986, Hartmann et al. 1986, 1990). Suppose that the partial's pitch and the probability of harmonic fusion both depend on an internal variable function of the partial's frequency. Suppose further that this variable is noisy (distributed randomly from trial to trial). When the variable falls near the harmonic series, fusion occurs and the pitch match fails. When it falls away from the series, the partial is easier to hear. The distribution of successful (unfused) pitch matches is therefore distorted, and its center of gravity shifted in the same direction as the mistuning.

3 Model

The internal random variable correlate of the partial's frequency is denoted as x. Its mean x_0 is proportional to the frequency f of the component, measured as a percentage mistuning from the harmonic series. Supposing that x is distributed normally around its mean x_0:

$$a(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-x_0)^2/2\sigma^2}$$

Supposing that the probability $b(x)$ of fusion conditional on x is shaped like a gaussian centered at zero mistuning:

$$b(x) = \alpha e^{-x^2/2s^2}$$

where s is the width of the "harmonic sieve", and α is a factor that determines the maximum probability of fusion. The distribution of successful matches is:

$$c(x) = A a(x)(1 - b(x))$$

where the normalization factor A ensures that the distribution sums to 1. Based on these assumptions, it is possible to calculate the pitch shift as the difference between the mean $\bar{x}$ of the distorted distribution $c(x)$ and the mean x_0 of the original distribution $a(x)$:

$$\bar{x} - x_0 = x_0 \frac{\sigma^2}{s^2+\sigma^2} / \left(\frac{\sqrt{s^2+\sigma^2}}{\alpha s} e^{x_0^2/2(s^2+\sigma^2)} - 1 \right)$$

The model has three parameters: σ, s and α.

4 Shifts produced by the model

Symbols in Fig. 1 represent shifts observed by Hartmann and Doty (1996) for the 5th harmonic at a low level (28 dB per harmonic). The line represents shifts produced by the model for $\alpha = 0.8$ and $s = \sigma = 3\%$ (these values were selected to give a good fit "by eye"). The predicted pitches are close to those observed experimentally. They have the same sign as

*xx
xx
xx

the mistuning, and peak at about 4% mistuning and decrease thereafter. With the same parameters the model can account for shifts observed at most of the other harmonics, both at 28 dB/harmonic and at 58 dB/harmonic. The somewhat larger shifts observed at harmonics 9 and 11 (28 dB/harmonic) and harmonic 7 (58 dB/harmonic) can be accomodated by assuming $s = \sigma = 4\%$. The monotonously increasing shifts observed at the fundamental (Fig. 2) can be accounted for by assuming $s = \sigma = 8\%$ and $\alpha = 0.6$.

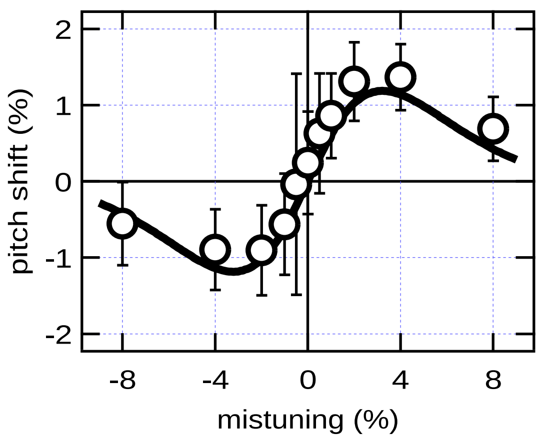

Fig. 1 *Symbols: pitch shifts observed for the 5th harmonic at a level of 28 dB per harmonic (Hartmann and Doty 1996). Line: shifts produced by the model assuming* $s = \sigma = 3\%$ *and* $\alpha = 0.8$.

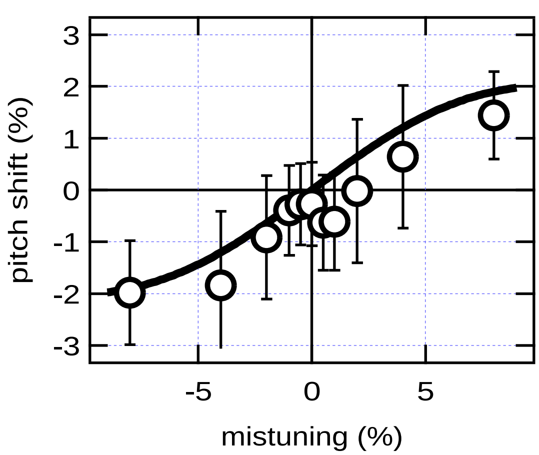

Fig. 2 *Symbols: pitch shifts observed for the fundamental at a level of 58 dB per harmonic (Hartmann and Doty 1996). Line: shifts produced by the model assuming* $s = \sigma = 8\%$ *and* $\alpha = 0.6$.

5 Discussion

The model accounts for the experimental pitch shifts quite well. The value (3 %) chosen for s and σ is consistent with the width of the "harmonic sieve" suggested by Moore et al. (1985, 1986), although a somewhat larger value was required to account for pitch shifts at some harmonics and/or levels. The fact that variability (σ), width of the harmonic sieve (s) or probability of fusion (α) differ between harmonics, levels or subjects is not particularly surprising. The role of harmonicity in the model is consistent with its role in harmonic sound segregation, in particular its effect on the number of sources perceived (de Cheveigné 1997a).

Similar shifts have been found for the pitch of tones preceded by a tone of similar frequency (Hartmann 1979), or the localization of sources preceded by a source of similar position (Kashino and Nishida, 1995). It is possible that a model similar to this one could account for those effects.

6 Acknowledgements

This work was carried out within a research agreement between ATR and CNRS. Bill Hartmann provided unpublished data and many useful comments and discussions.

References

[1] de Cheveigné, A. (1997a). “Concurrent vowel segregation III: A neural model of harmonic interference cancellation.” J. Acoust. Soc. Am., in press.

[2] de Cheveigné, A. (1997b)"Harmonic fusion and pitch shifts of mistuned partials", J. Acoust. Soc. Am. (submitted).

[3] Hartmann, W. M. (1979). "The effect of a prior tone on the pitch of a short tone," Proc. of the Research Symposium on the Psychology and Acoustics of Music, edited by W.V. May, 199-215.

[4] Hartmann, W. M., McAdams, S., and Smith, B. K. (1990). "Hearing a mistuned harmonic in an otherwise periodic complex tone," J. Acoust. Soc. Am. 88,1712-1724.

[5] Hartmann, W. M., and Doty, S. L. (1996). "On the pitches of the components of a complex tone," J. Acoust. Soc. Am. 99, 567-578.

[6] Kashino, M. and S. Nishida (1996). “Auditory localization aftereffects.” J. Acoust. Soc. Am. 99: 2596.

[7] Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1985). "Thresholds for the detection of inharmonicity in complex tones," J. Acoust. Soc. Am. 77, 1861-1867.

[8] Moore, B. C. J., Peters, R. W., and Glasberg, B. R. (1986). "Thresholds for hearing mistuned partials as separate tones in harmonic complexes," J. Acoust. Soc. Am. 80, 479-483.

[9] Terhardt, E. (1979). "Calculating virtual pitch," Hearing Research 1,155-182.