AN EFFECT OF THE COHERENCE BETWEEN ENVELOPES ACROSS FREQUENCY REGIONS ON THE PERCEPTION OF ROUGHNESS

DANIEL PRESSNITZER[1], STEPHEN MCADAMS[2,3]

(1) CNBH, Dept. of Physiology, University of Cambridge, U.K.

(2) IRCAM, Paris, France

(3) Laboratoire de Psychologie Expérimentale (CNRS), Université René Descartes, Paris, France

1 Introduction

1.1 Roughness and envelope coherence: within- and cross-channel cues

Auditory roughness arises when rapid beats can be perceived[1]. It is considered to be linked to the amount of amplitude fluctuation within a given auditory filter, around a "roughest" modulation frequency of 70 Hz[2].

The way roughness builds up if sounds have energy spread over a large frequency region is unclear. When two sinusoidally amplitude-modulated (SAM) tones with different carrier frequencies but a same modulation frequency are added, conditions with "co-phasic" envelopes are rougher than those with "anti-phasic" envelopes[2]. This could be accounted for by cross-channel comparisons of envelope coherence.

However, in places where the two SAM tones can interact along the basilar membrane, the resulting envelope in the anti-phasic conditions will be almost flat. The difference between co- and anti-phasic modulations could then be explained by within-channel cues only[3]. Some limited cross-channel processing had to be introduced in models in order to reproduce the low roughness of wide-band noise[3,4], but the hypothesised mechanism has not been directly tested. New stimuli were designed to address this issue.

The stimuli were derived from SAM-tones by introducing a time-jitter in their envelope on a period-to-period basis. A jAM tone can be described by its average modulation frequency $\overline{f_m}$, its centre frequency f_c, its amount of jitter (that controls the bandwidth), and its modulation depth m (that precisely sets the envelope rms value). All jAM tones with a same set of parameters should thus produce exactly the same roughness. It is however possible to obtain incoherent envelopes within a set of parameters by using different random samples for the jitter. Two jAM tones can then be added, without producing as important within-channel cues as anti-phasic SAM-tones.

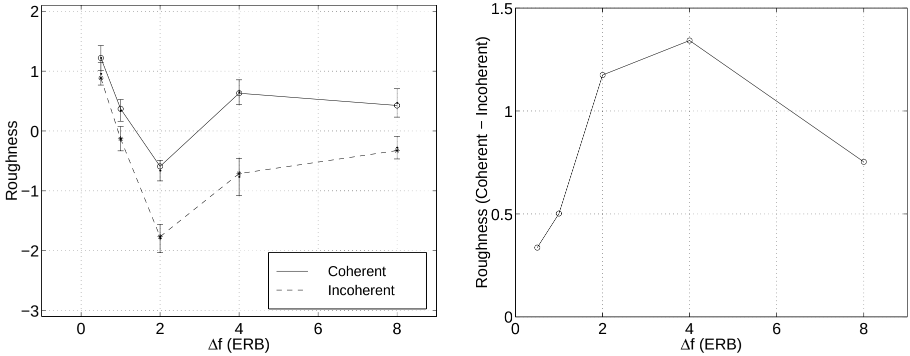

Figure 1. Experimental results. Left panel shows the relative roughness for the 12 stimuli. Right panel shows the coherence effect.

2 Experiment

2.1 Method

Each stimulus consists of the addition of two jAM tones with $\overline{f_m} = 50$ Hz and $m = 0.7$. The maximum amount of envelope period jitter was such that the range of possible frequencies for each period was between 33 and 100 Hz. One of the tones had a carrier of 4000 Hz and the second carrier took values between 4233 Hz and 9772 Hz. Values of the frequency difference between carriers were $\Delta_f = [1/2, 1, 2, 4, 8]$ on the ERB-rate scale. For half of the conditions, both jAM tones had the same envelope ("coherent"). For the other half of the conditions, two different random samples were used to compute the envelope jitter ("incoherent"). In this latter case, the correlation between envelopes was close to 0. Paired comparisons were used to collect roughness judgements. The experimental procedure and analysis methods have been described elsewhere[5]. Fifteen subjects took part in the experiment.

2.2 Results

The shape of the relative roughness values shown in Figure 1 is rather complex. For the purpose of the present paper, we will only focus on the "coherence effect" (the roughness difference between coherent and incoherent conditions). Overall,coherent envelopes yielded more roughness than the incoherent ones. This effect is low for a frequency separation of $\Delta_f = 1/2$ ERB, increases to a maximum at 4 ERB and then decreases but is still present at 8 ERB.

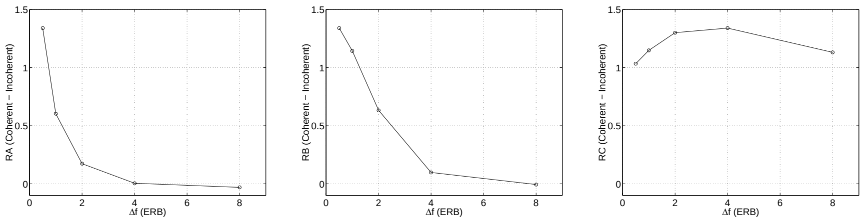

Figure 2. A comparison between the coherence effect predicted by the three roughness combination methods RA, RB, and RC described in the text.

3 Comparison of computational summation methods

Let us consider the way classic models of roughness would handle the stimuli. The signals would first be decomposed into i frequency channels and partial roughnesses r_i would be computed[2,3,4]. The next step is then to combine the partial roughnesses.

A straight-forward method is to add up the partial roughnesses across channels: $RA = \frac{1}{N}\sum_{i=1}^{N} r_i$. Although this method is surely wrong (because of the excessive roughness it would predict for white noise), it describes the overall influence of within-channel cues. A second method takes into account the correlation between the envelopes on which the roughness estimation is based[3,4]. Roughnesses in adjacent channels only add up if their envelopes are correlated: $RB = \frac{1}{N}\sum_{i=2}^{N-1}(c_{i,i-1}c_{i,i+1})^2 r_i$. Finally, all possible pairs of channels can be considered and roughness in any two channels only add up if the corresponding envelopes are correlated: $RC = \frac{1}{N(N-1)}\sum_{\substack{i,j=1 \\ j\geq i}}^{N} c_{i,j}(r_i+r_j)$.

The prediction of these algorithms are displayed in Figure 2, using a common front end to compute the partial roughnesses. The method RA only predicts a coherence effect if the excitation pattern of the sounds overlap. This is in clear opposition with the experimental results. For the same reasons, the method RB also fails, as it predicts an effect only if adjacent channels are involved. The last method is the only one to produce the correct trend, although the effect is systematically over-estimated to a substantial degree. It predicts a coherence effect proportional to the number of incoherent channels. As $\Delta_f$increases, this number first increases as the two jAM tones become separately resolved, and the coherence effect gets bigger. When the jAM tones do not overlap anymore, the number of incoherent channels decreases as the auditory filters widen in the frequency region of the higher jAM tone, and the coherence effect diminishes. This is in line with the experiment.

4 Discussion

The roughness of complex sounds not only depends on the partial roughnesses in different frequency regions but also on the coherence between the modulations present in these regions: incoherent modulations are less efficient than coherent ones. In the experiment we presented, this effect was clearly the result of cross-channel cues as it was small for small Δ_f where potential within-channel cues were maximal, and remained significant for large Δ_f where within-channel cues would be negligible.

Computational simulations indicated that a simple correlation between envelopes across all channels could qualitatively account for the trend in the results. However, we do not wish to imply that the auditory system performs such a correlation. Firstly, the same kind of effect could be achieved by some kind of cancellation across channels. Secondly, the coherence effect observed for a complex-sound attribute such as roughness could reflect the simultaneous action of cross-channel mechanisms of various kinds[6].

Finally, we only focused on one aspect of the data in this paper. There are actually other questions raised by the complex evolution of roughness with Δ_f and by measures made at $\Delta_f = 0$. Some of these issues are currently being investigated and they point to the need for a revision of the current "within-channel" roughness models before proceeding any further in trying to quantitatively adjust the cross-channel combination algorithm. [Work supported by the Fyssen Foundation.]

References

1. Hermann L. F. von Helmholtz. *On the Sensations of Tone as the Physiological Basis for the Theory of Music*. 1877. 2nd. Ed. trans. A. J. Ellis from German 4th Ed., 1885, Dover, 1954.
2. E. Terhardt. On the perception of periodic sound fluctuation (roughness). *Acustica*, 30:201–212, 1974.
3. W. Aures. Ein Berechnungsverfahren der Rauhigkeit [A roughness calculation method]. *Acustica*, 58:268–281, 1985.
4. P. Daniel and R. Weber. Psychoacoustical roughness: Implementation of an optimized model. *Acustica*, 83:113–123, 1997.
5. D. Pressnitzer and S. McAdams. Phase effects in roughness perception. In A. Rees, A. Q. Summerfield, and R. Meddis, editors, *Psychophysical and Physiological Advances in Hearing*, Whurr Publishers, 1998.
6. J. W. III Hall, J. H. Grose, and L. Mendoza. Across-channel processes in masking. In B. C. J. Moore, editor, *Hearing*, Academic Press, 1995.