Minimum differences of level and frequency for perceptual fission of tone sequences ABAB
Noorden, van, L.P.A.S.

Published in:
Journal of the Acoustical Society of America

DOI:
10.1121/1.381388

Published: 01/01/1977

Document Version
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:
• A submitted manuscript is the author's version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

Citation for published version (APA):
Minimum differences of level and frequency for perceptual fission of tone sequences ABAB*

Leon P. A. S. van Noorden

Institute for Perception Research, Den Dolech, Eindhoven, The Netherlands

(Received 7 June 1976; revised 22 November 1976)

Stream segregation or fission of the fast alternating tone sequence ABAB is known to occur if there is a sufficient frequency difference between the tones A and B. In this paper it will be shown that level difference instead of frequency difference can be sufficient to enable the occurrence of fission. The smallest level difference between A and B, $\Delta L \leq 3$ dB (2.5–10 tones per sec; tone duration 40 msec). At rates faster than 12 tones per sec a new perceptive phenomenon was observed: the roll effect. It is characterized by the weak tones being heard at double the tempo. The relation with the continuity effect is investigated using alternating sequences with both level and frequency difference between the tones as stimuli.

PACS numbers: 43.66.Lj, 43.66.Hg, 43.66.Mk, 43.66.Cb

INTRODUCTION

The alternating tone sequence ABAB... of two pure tones of different frequencies can split up perceptually into two simultaneous running sequences A.A. and B.B. (Miller and Heise, 1950; Dowling, 1968; Bregman and Campbell, 1971). The fission phenomenon occurs predominantly in fast tone sequences with large frequency separations. The attentional set of the observer also has a large influence on the occurrence of fission (Van Noorden, 1975). This leads to a distinction between the temporal coherence boundary, i.e., the largest frequency interval between $f_A$ and $f_B$ where the observer can still hear the alternation ABAB... and the fission boundary, i.e., the smallest frequency interval where the sequences A.A. or B.B. can be heard separately. The temporal coherence boundary depends heavily upon the tone rate, its value increases from about 3 semitones at a rate of 10 tones per sec to 15 semitones at 5 tones per sec. The fission boundary, however, is relatively independent of the tone rate; approximately one semitone over a large range of tone rates. This constancy led us to think about a possible relation between the fission boundary and the peripheral-frequency selectivity of the ear. The close relation of the “trill threshold” of Miller and Heise and the critical band (Licklider, 1951) is also a hint in this direction. A simple model of this relation would be a filter bank as a primitive description of the peripheral-frequency analysis, followed by a switch which removes the output of any of the filters, it can be selected.

It needs to be determined whether the selection process really takes place at such a peripheral level or, at a higher level, where the sounds are sorted with respect to features such as pitch, duration, and loudness. To discern the level of processing, one should study the phenomena in alternating tone sequences in which the tones do not differ by frequency but by another aspect, such as their amplitude. If it still turns out to be possible to select the strings of the A tones and the B tones we can reject at least the simplest model of peripheral selection (i.e., a filter set with an attentional switch). Tones of identical frequencies would pass through the same filter no matter what their amplitude.

In this article we want to report that we found it quite well possible to perceive fission with only an amplitude difference between the tones. Under the proper conditions one is not only able to hear the string of loud tones but the string of weak tones equally well. We consider this an interesting finding. It may give an answer to our question and it seems also to contradict our common intuition about loudness differences. A loud sound tends to mask a weaker one but one is not likely to think that a loudness difference can also help to distinguish different sound sources. In our case, however, the masking effect was eliminated because we used long enough silent pauses between the tones so that each tone pulse can be perceived as a distinct event.

Only a few references are made to this phenomenon. Dowling (1968) reports that interleaved melodies with overlapping frequency ranges can be recognized if there is an amplitude difference between the tones of the two melodies. Egan et al. (1959) found the same effect in relation to speech perception.

To make a reconnaissance of this phenomenon we measured the minimum level difference between A and B for perceptual fission of sequence ABAB... We call this the fission boundary, to be analogous with the case of fission by frequency difference. We varied the time between the tone pulses to study the extent of the fission boundaries for level and frequency differences. At very short times between the tones one might see the influence of the masking effects of the louder on the weaker tone. In fact, we discovered a new phenomenon in this region (See Sec. II).

I. FISSION BOUNDARY

A. Method

The stimulus is the monotone sequence ABAB... of pure tones A and B, $f_A = f_B = 1$ kHz, the level of tones B ($L_B$) is fixed at 35 dB SL while the level of tones A ($L_A$) is variable. The tone duration is 40 msec and the envelope trapezoidal with flanks of 5 msec. The tone repetition time $T$ used varies between 43 and 800 msec. Presentation is diotic via Sennheiser HD 414 headphones, while the subject is seated in a sound-proof booth.

The observer has to adjust the level of the A tones.

---


PACS numbers: 43.66.Lj, 43.66.Hg, 43.66.Mk, 43.66.Cb.

J. Acoust. Soc. Am., Vol. 61, No. 4, April 1977

Copyright © 1977 by the Acoustical Society of America.
The adjustment attenuator has a range of 20 dB in steps of 1 dB and is specially modified to give no audible and no tangible clicks. The position of the blinded knob is read out digitally, when the observer depresses a pushbutton. The results of the adjustments were not fed back to the observer and the measurements were repeated a number of times in different order of presentation.

The observer was instructed to listen for the string B of constant loudness and to make the difference in level between the A and B tones so small that he could just hear this string separately. After three adjustments with $L_A > L_B$, three adjustments were made with $L_A < L_B$ or vice versa. These six adjustments were all made with the same value of $T$. All 10 values of $T$ were used during a single session, in random order. The two normally hearing observers completed five sessions in three days. One observer, the author, had ample training in performing psycho-acoustical measurements and was an amateur musician. The other lacked these qualifications. The equipment has been described in detail elsewhere (Van Noorden, 1975).

B. Results and discussion
The results are presented in Fig. 1. We may distinguish three ranges of $T$: small, medium and large, with dividing points at about 100 and 400 msec. The smallest values of the fission boundary area found in the medium range. Here the fission boundary is more or less independent of $T$ and has a value of about 2–4 dB. There is no difference between the situations with $L_A > L_B$ and $L_A < L_B$. In the range of large $T$ values the fission boundary increases with increasing $T$. The boundary remains symmetrical with respect to the sign of the level difference. In the range of the small-$T$ values, however, this symmetry is lost. When $L_A < L_B$ string B can be heard separately down to the shortest $T$ value employed (43 msec). The value of the fission boundary has a maximum of about 5 dB at 80 msec. When $L_A > L_B$ on the other hand, the value of the fission boundary increases sharply with decreasing $T$, and it is impossible to adjust the fission boundary at the measuring points $T=62$ and 43 msec.

There are striking similarities in the patterns for the fission boundary for frequency difference (Van Noorden, 1975) and for level difference. In both cases the fission boundary is independent of $T$ in the medium range of $T$ values. When we compare the value of the fission boundary in this region for frequency and level differences with the just-noticeable frequency and amplitude modulation, respectively (e.g., Zwicker, 1952), we find that in both cases the value of the fission boundary is roughly the same factor above the just-noticeable modulation (roughly a factor of 3 at 8-Hz modulation frequency). Further, the value of the fission boundary in both cases starts to increase with increasing $T$ at about $T=400$ msec. Memory constraints may be considered responsible. This similarity in the way the auditory system deals with frequency and amplitude differences leads us to reject a close relationship between the fission phenomenon and the peripheral frequency analysis system in the regions of medium and large $T$ values.

II. ROLL EFFECT
As we saw in the previous section, it proved impossible to hear string B at $T=43$ and 62 msec when the A tones were louder than the B tones, but the string A could be perceived in these conditions. If the observer directs his attention to the weak tones, he hears a string with the tempo of the string ABAB, i.e., twice the tempo of B tones alone, for moderate level differences. The tones of this fast string of weak tones seem to be of uniform loudness. It is as if the A tones are split into two parts, one part that can be heard separately as the string of the loud tones, and one that appears to be as weak as the B tones, contributing to the string of the double tempo. We call this effect the roll effect. (See Fig. 2.)

The roll effect can only be observed at moderate level differences. At smaller level differences the
string ABAB...will be heard. At larger level differences the soft B tones sound like a continuous tone. This is a well-known effect, which has been given, among others, the name, "continuity effect" (Houtgast 1972, 1973).

To find an explanation of the roll effect we consider the following simple model of the peripheral-frequency-analysis system. A pure tone gives rise to an excitation pattern on the basilar membrane. The maximum excitation is reached at a certain place determined by the frequency of the tone. Hair cells transmit the excitation to the neurons of the auditory nerve. The higher the level of the tone the broader the region in which the hair cells are stimulated; so that, if we have tones of the same frequency, the excitation region of a low-level tone lies completely within the excitation pattern of a high-level tone. We now make the assumption that the observer's selective attention acts by selecting neurons that originate at certain regions of the basilar membrane. It follows that it should be possible to select the string of loud tones by disregarding the neurons that carry signals from the weak tones. There are neurons which carry signals from the loud tones only. (See Fig. 3.) On the other hand, no neurons can be found that carry signals of the weak tones only, which should make it impossible to select the string of the weak tones. This is in agreement with the phenomena at the small values of T.

Further, these considerations lead to the prediction that the roll effect is not restricted to sequences in which tones A and B are of identical frequency. At a certain level difference the weaker tones may be shifted away from the frequency of the louder tones over a limited distance without changing the condition that the range of excitation of the weaker tones completely falls within the range of excitation of the louder tones. From the shape of the excitation patterns it follows that the distance over which the tones may be shifted without losing the roll effect should increase with the level difference.

In the next experiment this prediction is verified in the following way. The stimulus is the alternating tone sequence ABAB. The observer directs his attention, to the sequence of tones B which are fixed in frequency and level. He has control over the frequency of the tones A. The experimenter sets the level of tone A at a constant level difference from tone B. The observer starts with a large frequency difference so that he can hear clearly the separate sequence of tones B. Next he decreases the frequency difference slowly until he just hears the tones B at a faster rate (roll effect). This we called the roll threshold. To find the changeover between the phenomena at small-T values and those at medium-T values, measurements were made at several T values. (At the medium-T values the observer does not adjust the roll threshold but the fission boundary. In both cases, however, the adjustment is the point where he ceases to hear the percept of the separate sequence BB.)

A. Method

The same apparatus and the same stimulus are used as before, except that now \( f_A \neq f_B \). The B tones are fixed at 1000 Hz and at 35 dB SL. The level of the A tones is set by the experimenter at one of seven values between \( L_A - 5 \) and \( L_A + 30 \) dB. At a given value of \( L_A \), the observer makes three adjustments with \( f_A < f_B \), and three with \( f_A > f_B \). All values of \( L_A \) are dealt with in a random order at a constant value of T in a single session. Each observer had four sessions for each value of \( T \) (48, 62, 72, 81, and 100 msec, respectively).

B. Results and discussion

It can be seen from Fig. 4 that, at small values of T, the adjustment results in an open "V" curve and, at larger-T values in a closed "O" curve (at levels \( \Delta L = 10 \) and \( -5 \) dB for \( T = 100 \) msec both observed stated that fission occurs again at all frequencies \( f_A \). The transition between V and O curves is gradual. There is a slight discrepancy with the results of Sec. I above where the observers still adjusted a value of the fission boundary at \( T = 81 \) msec. Now, however, the O-shaped

![FIG. 2. The difference between fission and roll. In the case of fission in monotonic sequences with a level difference between the alternating the observer is able to hear either the string of loud A tones or the string of weak B tones, at will. In the case of roll the observer is able to hear the string of loud A tones but not the string of weak B tones. When he directs his attention to the latter he perceives a string of weak tones with twice the tempo of string B.](image-url)

![FIG. 3. Excitation patterns of two tones along the basilar membrane, differing in frequency and level. As long as the apex of A lies within the dotted V-shaped curve the excitation pattern of the high-level A tones will overlap the excitation pattern of the low-level B tones.](image-url)
The roll threshold and fission boundary for the alternating sequence ABAB... with frequency and level differences between A and B. The observers adjusted $f_A$ so that they could just hear the string of tones B separately. Each experimental point is the mean of the results for two observers, who made 12 adjustments each. As can be seen, the roll-threshold curves are V-shaped and the fission boundary is a closed O-shaped curve. As $T$ increases the V-shaped curves fold up to form the O-shaped curve when $T$ equals 100 msec. The V curves at the shortest-$T$ values are in line with our expectations that the region in which the roll threshold may be observed broadens as the level difference increases; this V curve can be considered as a portrayal of the excitation pattern of the loud tones at a peripheral level. To check whether the slopes are in agreement with the slopes of curves that reflect the peripheral neural excitation pattern we measured the pulsation threshold (Houtgast, 1973) in the same tone sequence. Only the adjustment criterion is changed in these measurements. At a certain level difference between the A and B tones set by the experimenter, the observer had to adjust the frequency $f_A$ as far as possible from the frequency of the B tones so that he could still only just hear the tones B sound like a continuous tone (pulsation threshold). Since the pulsation threshold depends to a large extent on the duration of the silent interval between the tones, it could only be determined at $T = 48$ and 62 msec.

The results are plotted in Fig. 5 together with roll thresholds obtained earlier at the same values of $T$. The measured pulsation thresholds are in agreement with the measurements of Houtgast (1973) as regards the asymmetry and with those of Verschuur (1974, personal communication), who has shown that the pulsation threshold shifts "upwards" when there are small silent intervals or, in other words, that the larger the silent interval between the tones, the higher the level of the loud tones must be made in order to produce continuity. It can be seen in Fig. 5 that the pulsation threshold shifts upwards by about 20 dB as $T$ increases from 48 to 62 msec.

From the comparison of the two thresholds it is clear that the roll threshold mimics the pulsation threshold at the shortest value of $T$. The slopes of the two thresholds are parallel. The dependence upon $T$, however, is different in both cases. As we have seen above, the roll threshold folds gradually together with increasing $T$ to end in the O-shaped curve at about $T = 100$ msec. The pulsation threshold shifts upward keeping more or less a V shape. It may be better, however, to state that the pulsation threshold depends upon the silent gap between the tones, as follows from the fact that continuity can be observed in much slower tone sequences if the tones are longer (Houtgast, 1974). This is not the case with the roll threshold. At value of $T$ above about 80 msec we could not observe the roll no matter what the tone duration was.
Houtgast (1973) considers that the pulsation threshold reflects a neural excitation pattern. The fact that the continuity effect can also be observed when there are short pauses between the tone bursts indicates that this excitation does not cease immediately after the tone stops, but decays gradually (cf. Plomp, 1964). As long as the excitation at the start of the weak tone in an alternating sequence does not rise above the decaying excitation of the loud tone, the soft tone will sound continuous.

The form of the roll-threshold curve suggests that this threshold is also a reflection of the same neural excitation pattern; the relative positions of the roll and the pulsation thresholds indicate that the roll effect is produced when the excitation at the start of the weak tones is sufficiently far above the decaying excitation of the loud tones. The fact that we cannot perceive the soft tones as a separate string as soon as the excitation of the soft tones exceeds the decaying excitation from the loud tones, which is the case with somewhat larger tone repetition times, indicates that time is needed to process a tone completely up to the level at which the tones can be distinguished in loudness and frequency. Once this time is allowed for, the observer can select the loud tones and the weak tones equally well and hear the strings of these tones. But if the tones follow each other too rapidly he has to move the attentional constraints to a more peripheral level of perceptual processing.

**ACKNOWLEDGMENTS**

I thank my colleagues from the Institute for Perception Research, Eindhoven, The Netherlands, for support and discussion. Special thanks are due to B. L. Cardozo for guidance and H. van Leeuwen for skillful technical assistance.

*This paper covers a part of the thesis, "Temporal Coherence in the Perception of Tone Sequences," submitted to the Eindhoven University of Technology, The Netherlands, in February 1975 under the supervision of Professor Dr. J. F. Schouten. This research was supported by a grant from the Netherlands Organization for the Advancement of Pure Research (Z. W. O.).

In contrast to the case of frequency differences (Van Noorden, 1975) we do not know any application of this effect in music. It is clear that sequences of identical tones or beats which differ in intensity occur in music (e.g., to create rhythmic patterns) but these do not split up into two concurrent patterns. This effect would occur most clearly in fast sequences with large level differences between successive tones. Indeed, if one listens with headphones, it is possible to find settings of the level difference and the tone repetition time in the alternating tone sequence ABAB... at which there is inevitable fission (so there also has to be a temporal coherence boundary in these monotone sequences with level differences), but we could not find this phenomenon when listening with loudspeakers in a normal room. The weak tones were not audible at all. Obviously the reverberation time of the room hinders the clear, unambiguous, occurrence of fission. This was found when we tried to compose demonstrations of the phenomenon for the gramophone record which was included with the thesis and can be obtained separately from the Institute for Perception Research, P. O. Box 513, Eindhoven, The Netherlands.


