Perception & Psychophysics
1981, 29 (4), 323-335

The perceptual onset of musical tones

JOOS VOS
Institute for Perception TNO, Kampweg 5, Soesterberg, The Netherlands

and

RUDOLF RASCH
Institute of Musicology, University of Utrecht, Utrecht, The Netherlands

The perceptual onset of a musical tone can be defined as the moment in time at which the stimulus is first perceived. In the present experiments, a simple threshold model for the perceptual onset was applied. A paradigm was used in which a sequence of tones had to be adjusted in such a way that the onsets were perceived at equally spaced moments in time. In Experiment 1, the threshold model was applied in a design in which the rise times of the tones were varied. We concluded that the perceptual onsets of the tones can, indeed, be defined as the times at which the envelopes pass a relative threshold of 15 dB below the maximum level of the tones (82 dB). In Experiment 2, the maximum levels of the tones were varied from 37 to 77 dB. The results show that there is a shift in the relative threshold, but that this shift is small relative to the shift in the stimulus level. In Experiment 3, the effect of level above masked threshold on the perceptual onset was investigated in more detail by varying the level of a background noise. The results show that the relative threshold decreases with increasing level above masked threshold. The results from our experiments strongly suggest that the relative threshold is linearly dependent on the level above masked or absolute threshold and that a 7-dB increment of this level results in a 1-dB relative threshold decrement. The threshold model is compared with a current temporal integration model for the perceptual onset of tones. It is shown that our data cannot be adequately explained by temporal integration. Our experimental results suggest that adaptation of the hearing mechanism to a certain relative stimulus level is responsible for perceptual onset. The applicability of our threshold model in various realistic musical situations is discussed.

The perceptual onset of an acoustic stimulus, such as a musical tone or a speech syllable, can be defined as the moment in time at which the stimulus is first perceived. The physical onset, however, can be defined as the moment at which the generation of the stimulus has started. Generally, the perceptual onset is delayed relative to the physical onset. The time difference between physical and perceptual onset is caused by, among other things, the fact that most music and speech stimuli are not immediately at their maximum level, but begin with a gradually increasing amplitude. At the very beginning of the physical stimulus, its level is too low to attract the conscious attention of the listener. Only after the amplitude has increased to a certain level does the listener become aware of the presence of the stimulus. This first portion of an acoustic stimulus is called the rise portion. It is temporally defined either as a time interval between the physical onset and the moment that the maximum level is reached (this definition is used for percussive sounds) or as the time interval between the physical onset and the moment that a level of 3 dB below maximum level is reached (this definition is used for nonpercussive sounds). The durations of rise portions are within the range of 5 to 100 msec in most cases and depend on the kind of instrument, on the frequency of the tone played, and on the way in which the player starts the tone (Luce & Clark, 1965; Melka, 1970).

This paper is an extended written version of a contribution to the Third Workshop on Physical and Neuropsychological Foundations of Music, Ossiach/Austria, August 8-12, 1980. The authors are indebted to Reinier Plomp, Ino Flores d'Arcais, Andries Sanders, and Lex van der Heijden, who made useful comments on an earlier version of this paper. Thanks are also due Theo van Veen, who was very helpful with calculations on the convolution integrals, and Bert Schouten, for copyediting the manuscript. Requests for reprints should be sent to Joos Vos, Institute for Perception TNO, Kampweg 5, Soesterberg, The Netherlands.

Adequate description of the perceptual onset is very useful in psychoacoustical experiments designed to investigate the effects of temporal structure. Thus, in dichotic listening experiments, at least a certain amount of variance can be eliminated if the manipulation of the temporal order of the stimuli is expressed in perceptual onset asynchrony instead of physical onset asynchrony (see Marcus, 1976). The same possibly holds for diotic and dichotic temporal order judgments. Again, in the performance of ensemble music, perfect (subjective) synchronization is realized only when the perceptual onsets of the simultaneous tones in different voices coincide (Rasch,

0031-5117/81/040323-13$01.55/0

1979). Also, in research on the production of isochronous rhythm patterns, the musician isochronizes the perceptual onsets of the tones (Gabrielsson, 1974; Michon, 1967; Vos, 1973). The concept might also contribute to the understanding of prosody, an important factor in models of speech recognition. Finally, perceptual onset is probably a relevant parameter in the synthesis of music and speech by electronic devices.

During the last decade, a number of investigations relevant to the question of the perceptual onset of an acoustic stimulus have yielded some models that describe perceptual onset as a function of stimulus parameters.

In Schütte's model (1977, 1978a), perception is simulated by a first-order RC integrator circuit (leaky integrator), which is characterized by its time constant, τ. Inputs are tones with physical envelope functions, and outputs are subjective envelopes. Schütte defined the perceptual onset of a tone as the moment at which the subjective envelope passes a certain percentage of its maximum value. It should be emphasized that, in this model, a variable threshold, which depends on the physical envelope and tone duration, is used.

It should be noted, however, that Efron (1970a, 1970b, 1970c), in a coherent set of experiments on the relationship between the objective and subjective duration of a stimulus, found that perceptual onset was independent of stimulus duration.

In his 1970b experiments, Efron asked his subjects to report whether the onsets of two dichotically presented stimuli were simultaneous or not. The duration of the tone burst was fixed, the duration of the noise burst was varied, and the level of the noise burst was adjusted to maintain equal loudness. From experiments using another method (Efron, 1970c), as well as from experiments using cross-modal simultaneity judgments (Efron, 1970a), the same conclusion was drawn, that is, that stimulus duration is not relevant.

In the context of speech production and perception, the psychological moment of occurrence, termed the perceptual center (P-center) of syllables has been studied by Fowler (1979), Marcus (1976), and Morton, Marcus, and Frankish (1976). Morton et al. (1976) assumed that P-centers were a property of the acoustic make-up of the stimulus, although they failed to uncover a relevant marker for it.

Fowler (1979), however, questioned the significance of models that describe P-centers as a function of articulation-free acoustical parameters. Marcus (1976) described the P-center in a two-parameter finite-state model. This model involves an acoustic correlate of vowel onset (peak increment) as well as stimulus onset and offset. Stimulus onset and offset are defined as the time at which the temporal envelope of the signal intersects a threshold of about 30 dB. Peak increment and its associated time of occurrence were defined as the largest increment in spectral energy between consecutive time slices in a frequency band of 400 to 1,500 Hz.

A Simple Threshold Model Concerning the Perceptual Onset of Musical Tones

The experiments to be described in this paper were designed to apply a simple threshold model concerning the perceptual onset of musical tones. The physical temporal envelope of a musical tone can be roughly divided into three successive portions, the rise, the steady-state, and the decay portions (see Figure 1). As a function of time, the temporal envelope E(t) can be described as follows:

Rise portion:

$$E(t) = R\left(\frac{t - t_a}{\varrho}\right) \qquad t_a \leqslant t \leqslant t_a + \varrho$$

Steady-state portion:

$$E(t) = 1 \qquad t_a + \varrho \leqslant t \leqslant t_d = t_a + \delta \tag{1}$$

Decay portion:

$$E(t) = D\left(\frac{t - t_d}{\sigma}\right) \qquad t_d \leqslant t \leqslant t_d + \sigma,$$

in which E(t) = temporal envelope, R(t) = relative rise function, D(t) = relative decay function, t_a = physical onset (beginning of rise portion), t_d = physical offset (beginning of decay portion), ϱ = rise time (duration of rise portion), σ = decay time (duration of decay portion), and δ = tone duration (= $t_d - t_a$). The relative rise function describes the en-

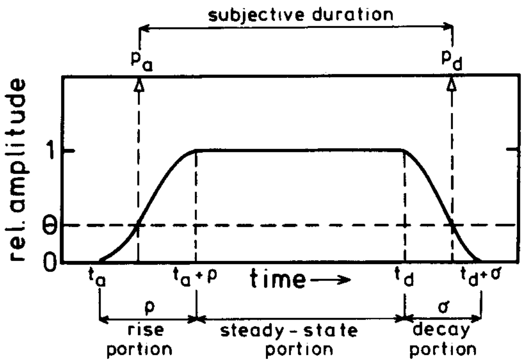

Figure 1. Temporal envelope of a musical tone, divided into three successive portions. The perceptual onset (p_a) and offset (p_d) of the tone are defined as the moment at which the temporal envelope passes relative threshold amplitude, Θ.

velope during the rise portion with the rise time ϱ as time unit and the maximum amplitude as amplitude unit. This means that it is defined only for $0 \leqslant t' \leqslant 1$ (t' equals $(t-t_a)/\varrho$). If the rise function $R(t) = v$ (v is relative amplitude) is monotonously increasing, then the inverse function $R^{-1}(v) = t$ is unambiguous. Throughout this paper, we will regard rise functions as monotonously increasing functions. The relative decay function is defined in the same way as the relative rise function. In our model, the perceptual onset of a tone p_a is the moment at which the temporal envelope during the rise portion passes a certain relative threshold amplitude Θ:

$$R\left(\frac{p_a - t_a}{\varrho}\right) = \Theta, \qquad (2)$$

in which p_a = perceptual onset and Θ = relative threshold amplitude. If p_a is known, Θ can be calculated with the help of Equation 2. If Θ is known, the perceptual onset can be calculated with the help of

$$R^{-1}(\Theta) = (p_a - t_a)/\varrho$$

or (3)

$$p_a = R^{-1}(\Theta) \cdot \varrho + t_a.$$

In the same vein, the perceptual offset, p_d, is defined as the moment at which the decay portion of the temporal envelope crosses the relative threshold amplitude, which may or may not be the same as the threshold amplitude of the perceptual onset. The subjective duration of a tone is defined as the time interval between the perceptual onset and offset of a tone. The following paragraphs, describing an extension of our model, will deal only with the perceptual onset.

The model can be extended to groups of tones, either simultaneous or successive ones. Two tones are called *perceptually synchronous* when their perceptual onsets coincide in time. Thus, for two tones, A and B, that are perceptually synchronous, as illustrated in Figure 2, with physical onsets t_a and t_b, rise functions R_a and R_b, rise times ϱ_a and ϱ_b, and maximum amplitude 1, the perceptual onsets p_a and p_b coincide, so that

$$p_a = R_a^{-1}(\Theta) \cdot \varrho_a + t_a = p_b = R_b^{-1}(\Theta) \cdot \varrho_b + t_b$$

or (4)

$$R_a^{-1}(\Theta) \cdot \varrho_a - R_b^{-1}(\Theta) \cdot \varrho_b = t_b - t_a.$$

This equation has only one unknown parameter, the relative threshold amplitude Θ. The other terms are known in an experimental situation as either independent or dependent variables. So it is possible to determine the relative threshold amplitude with the help of this equation. If the rise functions and rise times are identical ($R_a = R_b$, $\varrho_a = \varrho_b$), the perceptually synchronous condition results in $t_a = t_b$. In this case, the envelopes of the rise portions of the tones will coincide entirely, and every amplitude between 0 and 1 will be a solution of the above equation. The equation cannot be used, then, to determine the threshold amplitude. All this means that an experimental paradigm with simultaneous tones with different rise functions and/or rise times can be used to determine the threshold amplitude for perceptual onset.

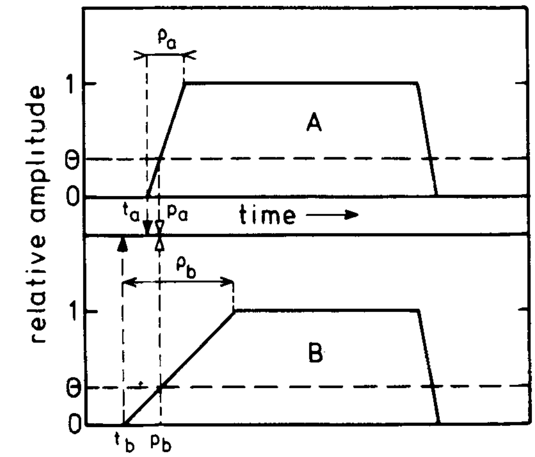

Figure 2. Temporal structure of a perceptually synchronous stimulus, containing tones A and B with physical onsets t_a and t_b, rise times ϱ_a and ϱ_b, and maximum amplitude 1. The perceptual onsets, p_a and p_b, coincide in time and are by definition located at the moment at which the temporal envelopes pass threshold amplitude, Θ.

If two tones have coinciding physical onsets ($t_a = t_b$), but different rise functions and/or rise times, the perceptual onsets will not, as a rule, coincide. Tone sequences are defined as *perceptually isochronous* if the time intervals between successive perceptual onsets are all equal to each other: $p_{n+1} - p_n = T$, for all values of n, in which T is the time interval between successive perceptual onsets. Figure 3c shows the temporal envelopes of the tones A, B, and A′, which are perceptually isochronous. If we express p_a and p_b in the inverse rise functions R_a^{-1} and R_b^{-1}, as was done above for perceptually synchronous tones, the following equation results:

$$R_a^{-1}(\Theta) \cdot \varrho_a - R_b^{-1}(\Theta) \cdot \varrho_b = T - (t_b - t_a). \qquad (5)$$

From this equation, we can solve the relative threshold amplitude Θ under the same conditions as were necessary for solving Equation 4. That means that an experimental paradigm with successive perceptually isochronous tones with different rise functions and/or rise times can also be used to determine the threshold amplitude for the perceptual onset.

If tones are physically isochronous, that is, when the time intervals between successive physical onsets are all equal to each other but have different rise times and/or functions, the perceptual onsets will not, as a rule, be isochronous.

In the model as described, there are only two stimulus variables, viz, rise time and rise function. The effect of rise time was investigated in the first experiment. Sensation level of the tones was varied in the second and third experiments and proved to have an effect on perceptual onset.

EXPERIMENT 1

Method

Procedure. We used a paradigm in which a sequence of tones had to be isochronized, that is, the tones had to be adjusted in such a way that the onsets were perceived isochronously. Each trial started with a tone sequence that was decidedly nonisochronous. The starting sequence consisted of successive pairs of tones, A and B, with a different interval between the physical onsets of A and the following B and between B and the following A (see Figure 3a). The onset times of tones B, relative to those of A, could be adjusted by the subjects, by turning a knob. The experimental task was to adjust the onset times of tone B in such a way that the sequence ABABAB . . . was perceptually isochronous, that is, that the perceived onsets of the tones followed each other with strictly the same time interval. This is illustrated in Figure 3b. Because the tones A were repeated every 800 msec, the subjective repetition time T of the tones in the entirely isochronized sequence is 400 msec. Rise times were varied independently. The time interval $t_b - t_a$ was derived from the position of the turning knob at which the subject judged the tone sequence to be isochronous. Now all variables that are necessary for computing the threshold amplitude for the perceptual onset from Equation 5 were known.

In our tone sequences, the physical offset times were kept isochronous, independent of the physical onset times. That means that tones B had different durations at the various stages of adjustment. However, since the perceptual offsets could also be considered perceptually isochronous, the subjective tone durations of tones A and B must have been equal in the final adjustments with isochronous perceptual onsets. Because of this relation between subjective onsets and durations, subjective duration could be (but not necessarily had to be) used as an extra cue for the isochronous adjustment.

An experimental series comprised 10 trials, which were all replications of the same condition. Four series were run consecutively. Each subject was tested individually. An experimental run lasted about ½ h. Between the runs there was a ½-h rest period, during which another subject was tested. The subjects were trained in the first runs. Knowledge of results was given to the subjects only with respect to the standard deviation of the 10 adjustments within a series. Standard deviations greater than 20 msec were exceptional. If they occurred, the series were repeated until results with standard deviations less than 20 msec were obtained.

The onset times of the tones B relative to those of tones A in case of a perceptually isochronous series were the all-important dependent variable to be determined. The relative amplitude of the threshold for the perceptual onset was calculated as follows. In our experimental conditions, the rise functions of tones A and B were equal ($R_a = R_b$), so that we may say: $R_a^{-1}(\Theta) = R_b^{-1}(\Theta) = \alpha$. We set $t_a = 0$ and $\Delta t = T - t_b$. Then Equation 5 can be simplified to

$$\alpha(\varrho_a - \varrho_b) = \Delta t$$

or (6)

$$\alpha = \Delta t/(\varrho_a - \varrho_b).$$

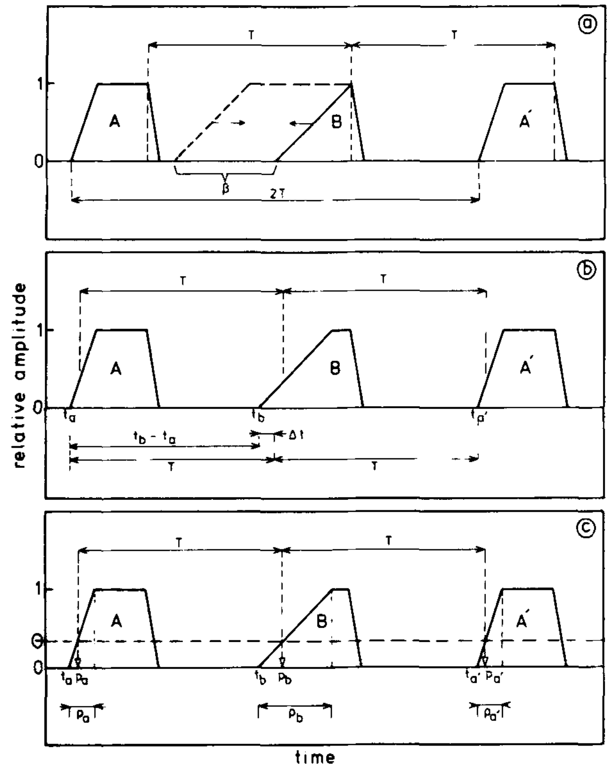

Figure 3. (a) Illustration of a perceptually nonisochronous starting sequence. The physical onsets of tones B could be adjusted by the subject within time interval β. At the start of a trial, the physical onset of tone B was either at the beginning or at the end of β. The physical onsets of tones A, as well as the physical offset times of tones A and B, were fixed. (b) Temporal structure of a perceptually isochronous tone sequence in which the time intervals between successive perceptual onsets are all equal to each other. The physical onsets of tones A and B are t_a and t_b, respectively. The repetition time of tone A equals 2T, so that T is the perceptual repetition time. The dependent variable is denoted by the time interval $\Delta t = T - (t_b - t_a)$. (c) Temporal structure of tones A, B, and A′ that are perceptually isochronous. The time interval between successive perceptual onsets is denoted by T. The perceptual onsets p_a, p_b, and $p_{a'}$ are defined as the moment at which the temporal envelopes pass relative threshold amplitude, Θ.

Since $\alpha = R^{-1}(\Theta)$, $R(\alpha) = \Theta$. Now, Θ represents the relative threshold amplitude. In our experiments, we used the relative rise function

$$R(t') = 0.5 + 0.5 \sin[(t' - 0.5)\pi].$$

Thus,

$$R(\alpha) = 0.5 + 0.5 \sin\left(\frac{\Delta t}{\varrho_a - \varrho_b} - 0.5\right)\pi. \tag{7}$$

In this formula, Δt is the dependent experimental variable. It can be computed from the repetition time T (=400 msec) and the adjusted physical onsets of tones B t_b.

The rise function and the rise times are determined as features of the experimental conditions. If tones A and B have equal rise times ($\varrho_a = \varrho_b$), Equation 6 cannot be solved. In this condition, the envelopes should coincide in the case of a perceptually isochronous tone sequence, and Δt should be zero.

Stimuli. Waveforms were calculated with the formula

$$p(t) = \sum_{n=1}^{n=20} (1/n) \sin 2\pi nft. \tag{8}$$

This results in a waveform with a spectral envelope with a slope of −6 dB/octave. Fundamental frequencies of tones A and B were 400 Hz. The duration of tones A was always 150 msec; the duration of tones B depended on the adjusted onset of the moment. Rise times of tones A and B were independently varied. Decay times of A and B were held constant at 20 msec. The absolute rise function, as shaped by the analog gates, is described by

$$R(t) = 0.5 + 0.5 \sin \{[(t/\varrho) - 0.5]\pi\}. \tag{9}$$

In this formula, the physical onset is set at $t = 0$ and the maximum relative amplitude (1) is reached at $t = \varrho$. The decay function is described in a similar way:

$$R(t) = 0.5 - 0.5 \sin \{[(t/\sigma) - 0.5]\pi\}. \tag{10}$$

Here, t equals 0 at the beginning of the decay portion. In this paper, rise and decay times are referred to as time interval ϱ', which indicates the time interval necessary for the rise curve to increase from 10% to 90% of the maximum amplitude. For a rise or decay time ϱ' as defined above, the time ϱ between zero and maximum amplitude is given by:

$$\varrho = \varrho'(\arcsin 1/\arcsin 0.8) = 1.69\varrho'. \tag{11}$$

The sound-pressure level of the tones was 82 dB(A) measured as continuous signals.

Apparatus. A flow diagram of the apparatus is shown in Figure 4. The experiments were run under the control of a PDP-11/10

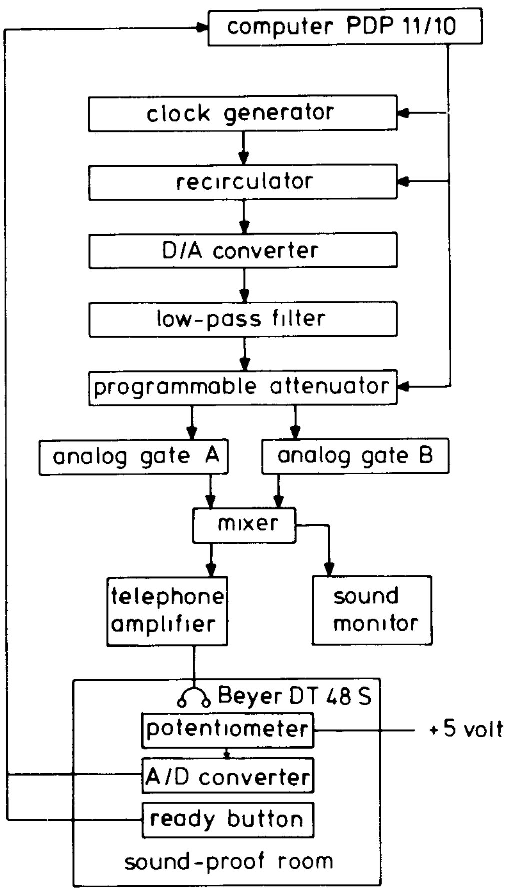

Figure 4. Flow diagram of the apparatus used. Tones A were established by means of gate A, tones B by means of gate B.

computer. A continuous tone was generated in the following way. One period of the waveform was stored in 256 discrete samples (with 10-bit accuracy) in an external revolving memory (recirculator), which could be read out by a digital-to-analog converter. The sampling rate was determined by a pulse train derived from a frequency generator. The tone was filtered by a low-pass filter with a cutoff frequency of 5 kHz. Sound-pressure level was controlled by a programmable attenuator. Tone A was presented by passing it through on/off gate A, tone B by passing it through on/off gate B. After gating, the tones were mixed and fed to a headphone amplifier. The signals were presented diotically (same signal to both ears) by means of headphones.

Subjects were seated in a soundproof room. The subjects had to find out first in which area of the blinded knob the onsets of tones B could be controlled. There were five such areas, which overlapped. Which area was sensitive was determined by a random procedure. Two consecutive trials could not have the same sensitive area. The voltage of the adjustment knob was read out by an analog-to-digital converter and was transformed into a time measure for the onsets of tones B relative to the onsets of tones A. Tests revealed that the accuracy of this measurement procedure, that is, the transformation of the voltage to the timing of t_b relative to t_a, was within 1 msec. When the subject considered the tone sequence to be isochronous, he pressed a ready-button.

Experimental design. The independent variables were: (1) rise time of tone A (5, 20, 40, 60, and 80 msec), and (2) rise time of tone B (5, 20, 40, 60, and 80 msec). In order to reduce the number of experimental conditions, an incomplete factorial design was chosen. The combinations in which rise times of tones B were shorter than the rise times of tones A were omitted. The 15 different conditions were presented twice to five subjects, so each mean value is based on 100 trials.

Subjects. The subjects were musically trained. Three of them were students from the Institute of Musicology at Utrecht. They were paid for their services. The two authors participated in the first experiment. Their data were not significantly different from those of the other (naive) subjects.

Results and Discussion

In Table 1, mean Δts are given for all presented combinations of the rise times of tones A and B. Analyses of variance (ANOVA) were carried out. Because of the characteristics of the design, the effects of the rise times of tones A and B were tested separately for each level of the A and B tone rise times. The ANOVAs showed a highly significantly effect of the rise time of tone B and of the rise time of tone A on Δt. F ratios and significance levels are represented in Table 2. Replications within and between series were not significantly different. The interaction between the rise times of tones A and B, testable for only part of the data [e.g., A (5, 20, 40) and B (40, 60, 80)], was not significant.

As shown above, in the conditions in which tones A and B have equal rise times, Δt should be zero. Inspection of the mean Δts for equal rise times in Table 1, however, reveals a small, but consistent, tendency to place the onset of tone B too early. This adjustment effect, the subject's bias towards turning the knob consistently too far in one direction, was also found by Marcus (1976), Schütte (1977), and Zwicker (1970). A similar effect, the occurrence of systematic time-order errors, was found when subjects were asked to produce a series of monosyllables in a rhythmical way (Fowler, 1979) and when auditory durations had to

Table 1
Mean Δts and Corrected Δts for All Experimental Combinations of Rise Times of Tones A and B, Together With the Corresponding Relative Amplitudes and Thresholds

Rise Time		Δt		Threshold for Perceptual Onset	
Tone A	Tone B	Observed	Corrected	Amplitude	Level
5	5	5.4	.5		
5	20	10.8	5.9	.128	−17.9
5	40	22.2	17.3	.197	−14.1
5	60	26.9	22.0	.132	−17.6
5	80	39.8	34.9	.176	−15.1
20	20	5.9	1.0		
20	40	13.2	8.3	.142	−17.0
20	60	24.6	19.7	.195	−14.2
20	80	34.4	29.5	.195	−14.2
40	40	5.7	.8		
40	60	15.1	10.2	.208	−13.6
40	80	25.0	20.1	.203	−13.9
60	60	3.3	−1.6		
60	80	13.6	8.7	.155	−16.2
80	80	4.2	− .7		

Note—Mean threshold level = −15.4 (SD = 1.6). It is not possible to derive thresholds for the conditions of equal rise times of tones A and B. The correction of Δt was −4.9 msec. Threshold level is given in decibels; Δts are given in milliseconds.

be compared or reproduced (Stott, 1935; Woodrow, 1934, 1935; Wundt, 1903; Vos, Note 1). The phenomenon seems to be independent of the task to be carried out.

Evidently, mean Δts are composed of a rise time effect (Δt_r) and an adjustment effect (Δt_a). It is assumed that Δt_a is not dependent on the rise times of tones A and B and that Δt is simply the sum of Δt_r and Δt_a. To test the assumption of the independence of the rise time effect and the adjustment effect, mean Δts from the conditions in which $\varrho_a < \varrho_b$ have to be compared with the mean Δts from the combinations in which $\varrho_a > \varrho_b$.

In our first experiment, we did not have combinations in which the rise time of tone A was longer than that of tone B. Therefore, we conducted another experiment in which all possible combinations of ϱ_a and ϱ_b were presented. Nine combinations (three values of the rise time of tone A and three values of the rise time of tone B) were presented to four other subjects. For the combinations in which ϱ_a is shorter or longer than ϱ_b, Δt is composed of Δt_r and Δt_a, while for the combinations in which ϱ_a equals ϱ_b, Δt equals Δt_a. If the rise time effect and the adjustment effect are mutually independent, the sum of the mean Δts from the conditions in which $\varrho_a < \varrho_b$ and the mean Δts from the conditions in which $\varrho_a > \varrho_b$ should equal twice the mean Δt of the conditions in which ϱ_a equals ϱ_b. (Note that Δt becomes negative when the physical onset of tone B is adjusted at $t > T$.)

In fact, this was confirmed by the data. We concluded that it was permissible to estimate Δt_a from the mean Δt of the conditions in which ϱ_a equals ϱ_b.

So the mean bias (4.9 msec) of the five conditions in our first experiment, in which ϱ_a equals ϱ_b, was subtracted from all Δts, and the corrected Δts (Δt_c) are given in the next column of Table 1. With the help of Equation 7, the relative threshold amplitudes (Θ) were computed with the Δt_cs for all combinations with unequal rise times. There were 10 different conditions in which the threshold level for the perceptual onset could be computed. The relative amplitudes and levels are presented in the third and fourth columns of Table 1. Mean threshold level is −15.4 dB, and the standard deviation equals 1.6 dB. Considering this consistent result—that is, that about the *same* relative threshold level was found in 10 physically different conditions—we are justified, for the time being, in defining the perceptual onset moment of a tone as the time at which its envelope passes a certain threshold level. Moreover, from our experimental data, we can estimate this threshold level as −15 dB, relative to the maximum level of the tones. In a recent study on synchronization in performed ensemble music, Rasch (1979) defined the perceptual onset of a musical tone as the moment that its envelope exceeded a threshold of about 15 to 20 dB below the maximum levels of the signals. So Rasch's adopted level is fairly compatible with our data.

In the model of Schütte (1977), the perceptual onset of a tone was defined as the moment at which the *subjective envelope* passed 16% of its maximum value. We found that the perceptual onset was located at the moment at which the *physical envelope* passed 17% of its maximum value. For the value of τ, adopted by Schütte, 16% of the *subjective* envelope *cannot* be equal to 17% of the *physical* envelope. We have to conclude that Schütte's model cannot explain our results. A more detailed comparison of our threshold model with Schütte's model will be made at the end of this paper.

Table 2
Results of the One-Way ANOVAs to Test the Effect of the Rise Time of Tone B for Each Magnitude of the Rise Time of Tone A Separately and the Effect of the Rise Time of Tone A for Each Magnitude of the Rise Time of Tone B Separately

Rise Time (in Milliseconds)				
Tone A	Tone B	df	F	p <
5	5,20,40,60,80	4,16	36.0	.00001
20	20,40,60,80	3,12	121.1	.000001
40	40,60,80	2,8	23.7	.001
60	60,80	1,4	46.2	.005
5,20,40,60,80	80	4,16	38.8	.000005
5,20,40,60	60	3,12	40.5	.00005
5,20,40	40	2,8	44.1	.0005
5,20	20	1,4	24.0	.01

EXPERIMENT 2

In Experiment 1, it was shown that the perceptual onsets of the alternating tones of equal intensity could be defined as the time at which the envelopes passed a threshold of −15 dB. The threshold was given relative to the maximum level of the tones, but it is evident that the threshold can also be defined as a level above background noise or hearing threshold. In short, the following question can be asked: Is the threshold fixed, with respect to maximum level, to background or to some other criterion? Our first experiment was designed to test the threshold hypothesis in general, not to discriminate between these alternatives. Experiment 2 was designed to determine the reference relative to which the threshold had to be defined. This was done by varying the maximum levels of the tones. Rise times were varied again, because they are a suitable tool for determination of the threshold of the perceptual onset. In addition, presentation of tones with different rise times is compatible with musical practice.

In a similar paradigm, Schütte (1977, 1978b) presented alternating tone pairs, one 20-msec tone and one 100-msec tone, in two different conditions in which the intensities of the tones A and B were held at a constant level. In the first condition, the levels of tones A and B were 70 dB. In the second condition, the tone pairs were presented at a very low signal-to-noise ratio: the levels of the tones were 3 dB higher than a continuous masker that had a level of 50 dB. The two conditions had about the same effect on Δt, which suggests that tone intensity has no effect on perceptual onset. Unfortunately, however, Schütte's results cannot solve our problem because his stimuli had rectangular envelopes. In terms of our model, such stimuli are not sensitive to determine a shift in the relative threshold.

Method

Procedure. The procedure was identical to that in Experiment 1.

Stimuli. The stimuli were the same as those in Experiment 1. The sound-pressure level of the tones, however, was 77 dB(A) in the highest level condition, and the tones were measured as continuous signals. The sound pressure levels of the tones in the other conditions were 57 and 37 dB(A).

Apparatus. The apparatus was the same as in Experiment 1.

Experimental design. The independent variables were: (1) rise time of tone A (5, 40, 80 msec); (2) rise time of tone B (5, 40, 80 msec); and (3) level of the tones [37, 57, 77 dB(A)]. Within a sequence, tones A and B had the same level. The 27 different conditions were presented in 54 experimental series to four subjects. Each mean value is based on 80 trials.

Subjects. Four students at the University of Utrecht served as subjects and were paid for their services. None of them had participated in the first experiment. They were musically trained.

Results

In Table 3, the corrected mean Δts of all experimental combinations of rise times of tones A and B for the three presented levels of tones A and B are given. The adjustment effects were computed for the 37-, 57-, and 77-dB intensity conditions separately, and they turned out to be 7.9, 8.1, and 7.2 msec, respectively.

An analysis of variance was carried out on Δt. For this purpose, a 4 (subjects) by 3 (rise times of tones A) by 3 (rise times of tones B) by 3 (tone levels) by 20 (replications) randomized block factorial design (Kirk, 1968) was used. The effects of the rise time of tone A [$F(2,6) = 1{,}258.1$, $p < .00005$] and the rise time of tone B [$F(2,6) = 612.8$, $p < .00005$] were highly significant. The significant interaction effect between the level and the rise time of tone A [$F(4,12) = 13.7$, $p < .0005$] showed that, with decreasing level, Δt increased when the rise time of tone A was short and decreased when the rise time of tone A was long. The interaction between level and the rise time of tone B revealed that, with decreasing level, Δt increased significantly when the rise time of

Table 3
Mean Corrected Δts of All Experimental Combinations of Rise Times of Tones A and B for the Three Presented Levels of Tones A and B, Together With the Corresponding Relative Thresholds

Rise Time (in Milliseconds)		Levels of Tones A and B: 37 dB (A) (Mean = −8.0, SD = .9)		57 dB (A) (Mean = −11.7, SD = .7)		77 dB (A) (Mean = −13.0, SD = 1.6)	
Tone A	Tone B	Δt_c	Threshold	Δt_c	Threshold	Δt_c	Threshold
5	5	− 3.5		− 4.1		− 2.2	
5	40	24.1	−9.0	21.2	−10.9	17.3	−14.1
5	80	58.1	−7.2	46.1	−10.7	41.2	−12.4
40	5	−28.1	−6.7	−19.5	−12.2	−21.1	−11.0
40	40	− .7		− 1.0		− 1.8	
40	80	30.4	−7.5	22.8	−11.9	24.0	−11.1
80	5	−53.3	−8.5	−42.5	−11.9	−36.5	−14.4
80	40	−27.5	−9.0	−21.8	−12.6	−18.8	−14.9
80	80	.9		− 1.4		− 2.2	

Note—Thresholds cannot be derived for the conditions of equal rise times of tones A and B. The correction of Δt was −7.9, −8.1, and −7.2 msec in the 37-, 57-, and 77-dB conditions, respectively. Threshold is given in decibels; Δt_c is given in milliseconds.

tone B was long and decreased when the rise time of tone B was short [$F(4,12) = 20.1, p < .0005$].

With the help of Equation 7, the relative threshold amplitudes were computed with the Δt_cs for all combinations with unequal rise times. The relative levels are presented in Table 3. In the 37-dB condition, the mean relative threshold was −8.0 dB and the standard deviation equaled .9 dB. In the 57- and 77-dB conditions, the mean thresholds were −11.7 and −13.0 dB, and the standard deviations were .7 and 1.6 dB, respectively.

Discussion

The results of Experiment 2 show (1) that the time difference between physical and perceptual onsets increases with decreasing tone intensity, (2) that this increase can be described as an upward shift in the relative threshold by which the perceptual onset is determined, and (3) that the shift in threshold is small relative to the shift in stimulus level. Therefore, the threshold can be most conveniently described relative to the maximum level of the stimulus.

EXPERIMENT 3

It is a matter of everyday experience that the audibility of sounds like speech and music may be decreased in the presence of other sounds. Thus an orchestra may partially mask the sound of a soloist. In Experiment 2, it was shown that the relative threshold for the perceptual onset slightly increased with reduction of the levels of the tones. This lowering of the maximum levels of the tones corresponds to a reduction of the sensation level. The purpose of the third experiment was to investigate the effect of signal-to-noise ratio on the relative threshold in more detail.

Method

Procedure. The procedure was identical to that in Experiments 1 and 2. The signal-to-noise ratio, defined here as a level above masked threshold, was varied by changing the level of a continuous noise. Signal level was held constant. In addition, at the beginning of the first run, the adequate levels of the continuous masker were determined individually for every subject, for three different signal-to-noise ratios of the tones. For a signal-to-noise ratio of 20 dB, for example, the level of the now isochronously presented tones was decreased by 20 dB and the level of the masker was adjusted until the tones could be detected in 50% of the cases.

Stimuli. The stimuli were the same as in Experiment 1. The sound-pressure level of the tones measured as continuous signals, however, was 77 dB(A). The continuous masker was pink noise with a spectral envelope slope of −3 dB/octave.

Apparatus. The apparatus was the same as in Experiments 1 and 2. In addition, after appropriate attenuation, the output of the noise generator was fed to the headphone amplifier.

Design. The independent variables were: (1) rise time of tone A (5, 40, 80 msec); (2) rise time of tone B (5, 40, 80 msec); (3) level above masked threshold of tones A and B (20, 30, 40 dB). Within a sequence, tones A and B had the same level. The 27 different conditions were presented in 27 experimental series to four subjects.

Subjects. Two new students from the Institute of Musicology at Utrecht served as subjects and were paid for their services. One of the authors (J.V.) and a colleague also participated in the experiment. All subjects were musically trained.

Results

In Table 4, the corrected mean Δts of all experimental combinations of the rise times of tones A and B for each of the three levels above masked threshold are given. The adjustment effects were computed for the 20-, 30-, and 40-dB conditions separately, and they turned out to be 8.5, 7.2, and 7.0 msec, respectively. An analysis of variance was carried out on Δt_c, because the adjustment effect was to some extent dependent on the conditions. The effects of the rise time of tone A [$F(2,6) = 112.0, p < .0001$] and the rise time of tone B [$F(2,6) = 128.0, p < .0001$] were highly significant. The significant interaction effect between level above masked threshold and rise time of tone A [$F(4,12) = 3.8, p < .03$] showed that, with decreasing level above masked threshold, Δt_c increased when the rise time of tone A was short and decreased when the rise time of tone A was long. The

Table 4
Mean Corrected Δts of All Experimental Combinations of Rise Times of Tones A and B for the Three Presented Levels Above Masked Threshold of Tones A and B, Together With the Corresponding Relative Thresholds for the Perceptual Onsets

Rise Time (in Milliseconds)		Level Above Masked Threshold of Tones A and B: 20 dB (Mean = −6.6, SD = .9)		30 dB (Mean = −7.1, SD = 1.2)		40 dB (Mean = −9.2, SD = 2.6)	
Tone A	Tone B	Δt_c	Threshold	Δt_c	Threshold	Δt_c	Threshold
5	5	− 3.8		− 3.0		− 1.3	
5	40	26.7	−7.5	26.8	−7.4	25.3	− 8.2
5	80	65.2	−5.6	62.8	−6.2	52.5	− 8.7
40	5	−31.3	−5.3	−31.5	−5.2	−31.0	− 5.4
40	40	− 2.0		− .7		− .8	
40	80	32.0	−6.8	29.5	−7.9	25.0	−10.4
80	5	−59.8	−6.8	−58.0	−7.3	−50.5	− 9.3
80	40	−29.8	−7.8	−28.2	−8.6	−21.0	−13.2
80	80	2.7		2.3		1.5	

Note—Thresholds cannot be derived for the conditions of equal rise times of tones A and B. The correction of Δt was −8.5, −7.2, and −7.0 msec in the 20-, 30-, and 40-dB conditions, respectively. Threshold is given in decibels; Δt_c is given in milliseconds.

interaction between level above masked threshold and rise time of tone B revealed that, with decreasing level, there was a trend for Δt_c to increase when the rise time of tone B was long and to decrease when the rise time of tone B was short [$F(4,12) = 2.3$, $p < .12$].

With the help of Equation 7, the relative threshold amplitudes were computed with the Δt_cs for all combinations with unequal rise times. The relative levels are presented in Table 4. In the 20-dB condition, the mean relative threshold was −6.6 dB and the standard deviation equaled .9 dB. In the 30- and 40-dB conditions, the mean thresholds were −7.1 and −9.2 dB and the standard deviations were 1.2 and 2.6 dB, respectively.

Discussion

From the results of Experiment 3 it can be concluded that the relative threshold for the perceptual onset of musical tones decreases with increasing level above masked threshold. It should be noted that tones A and B were held at a constant level of 77 dB(A). The results of Experiment 3 can be related to those of Experiment 2, in which a similar decrease of the relative threshold was found with increasing sensation level of the tones.

In Figure 5, mean relative thresholds from Experiments 1, 2, and 3 are plotted as a function of level above masked or absolute threshold. In Experiment 1,

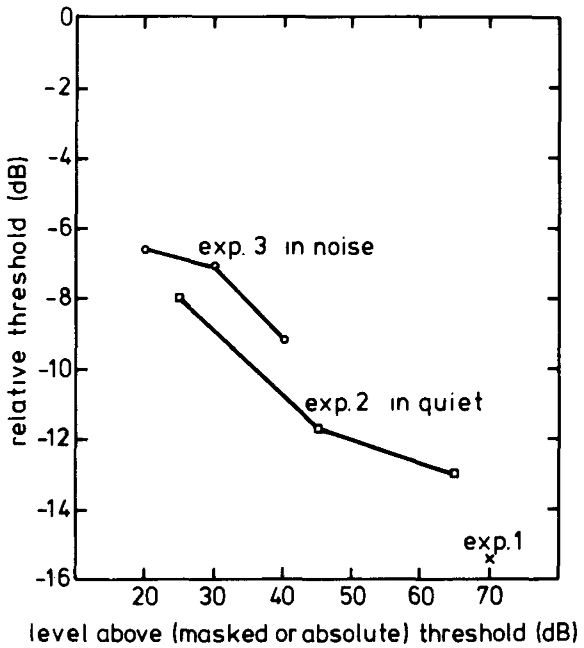

Figure 5. Mean relative threshold for the perceptual onset plotted as a function of the sensation level of tones A and B. For reference purposes, the mean relative thresholds from Experiments 1 and 2 are also plotted.

the stimulus level of 82 dB corresponded to a sensation level of 70 dB. In Experiment 2, the different tone levels of 37, 57, and 77 dB corresponded to sensation levels of 25, 45, and 65 dB, respectively. In Experiment 2, an increase of the sensation level of 40 dB resulted in a relative threshold decrement of 5 dB, while in Experiment 3, the increase of the level above masked threshold of 20 dB resulted in a relative threshold decrement of 2.6 dB. Moreover, the level of the relative threshold seems to be linearly dependent on the level above threshold. The relationship between level above threshold and relative threshold for perceptual onset can be roughly summarized by concluding that a 7-dB level above masked or absolute threshold increment results in a 1-dB relative threshold decrement. There remains a small difference between the levels of Experiments 2 and 3, for which no apparent explanation is available.

GENERAL DISCUSSION

From our experiments, we may conclude that the perceptual onsets of successively presented tones can be defined as the times at which the envelopes pass a *relative* threshold of about 6 to 15 dB below the maximum level of the tones. In a number of experiments, we have shown that the level of the relative threshold depends on the tone level above masked or absolute threshold. The data from the present experiments seem to suggest adaptation to a certain constant stimulus level. At the time the adaptation threshold is passed by the stimulus level presented, the onset of the stimulus is perceived. The experimental setup, in which sequences of alternating tones A and B with the same level were presented for a rather long time (one trial lasted about 30 to 120 sec), may have evoked optimal conditions for adaptation.

The nature of our threshold model is at variance with theories of temporal summation in hearing. In general, temporal summation theories suppose that the ear calculates a running average on the sound in accordance with the convolution integral:

$$y(t) = \int_0^t X(\tau)\, W(t - \tau)\, d\tau, \qquad (12)$$

where $y(t)$ is a central measure that forms the basis of the observer's response to the sound, $X(t)$ is the physical envelope of the stimulus, and $W(t)$ is a temporal weighting function. In the past three decades, a great many psycho-acoustic data both on the detectability of brief sounds and on loudness perception have been described by means of linear integration (see, e.g., Munson, 1947; Plomp, 1961; Plomp & Bouman, 1959; Zwislocki, 1960, 1969). Our threshold model will be compared with Schütte's (1978a) model for the perceptual onset of tones in more detail because Schütte's model is based on temporal

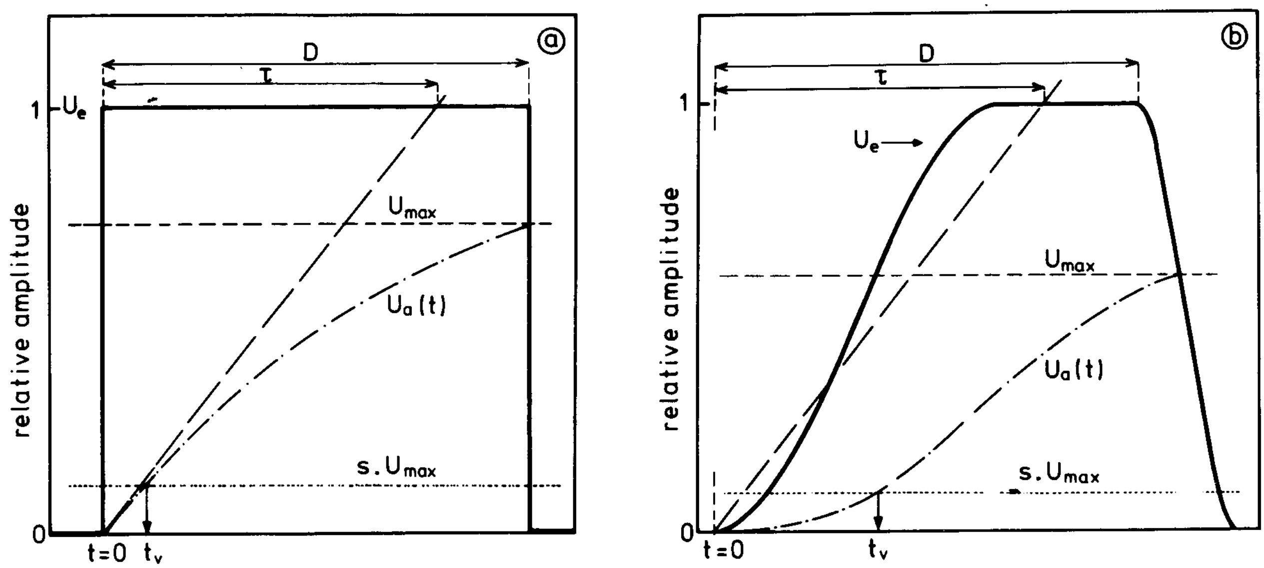

Figure 6. (a) Schematic diagram of the determination of the perceptual onset, after Schütte (1977). The solid line is the physical envelope, U_e. The duration of the stimulus is D. τ is the time constant of a low-pass filter that simulates perception. The dashed/dotted line is the subjective envelope, $U_a(t)$, which reaches its maximum level, U_{max}, at t = D. The perceptual onset (t_v) is located at the time at which $U_a(t)$ passes the relative amplitude threshold $s \cdot U_{max}$. (b) As in Figure 6a. The physical envelope, U_e, is an example of the envelopes used in Experiments 1, 2, and 3.

integration. First, we shall describe his model; second, we shall use the stimuli of our first experiment to compute the perceptual onset in terms of his model; and third, we shall test the predicting power of both models.

The Temporal Integration Model of Schütte

Schütte's (1977, 1978a) model is an extension of a model of Burghardt (1972) which describes the subjective duration of tones and pauses between tones. Perception is simulated by a first-order RC integrator circuit (leaky integrator), which is characterized by its time constant, τ. Inputs are tones with physical envelope functions, $U_e(t)$, and outputs are the subjective envelopes, $U_a(t)$. This is illustrated in Figure 6a. In the case of a rectangular physical envelope and tone duration, D ($U_e = 1$, for $0 < t < D$), the subjective envelope can be described by means of the following equation:

$$U_a(t) = U_e(t)[1 - \exp(-t/\tau)]. \quad (13)$$

As long as $U_e(t)$ is constant or increases, $U_a(t)$ increases monotonically. As a result of this, the maximum value of the subjective envelope, U_{max}, increases with tone duration D. The perceptual onset, t_v, of a tone is defined as the moment at which a threshold is passed. This threshold is defined as amplitude s, relative to U_{max}. For the stimulus depicted in Figure 6a, the perceptual onset t_v is determined by

$$t_v = -\tau \ln\{1 - s[1 - \exp(-D/\tau)]\}. \quad (14)$$

Schütte used the results of psychoacoustical experiments, in which pure tones had to be isochronized, to estimate τ and s in his model. The optimal values of τ and s were 120 msec and .16, respectively.

Computation of Perceptual Onset from Schütte's Model

The perceptual onsets of the tones that were presented in our first experiment were computed with the help of the convolution integral (Equation 12). In accordance with Schütte's model, the weighting function W(t) equaled $\exp(-\gamma t)$, where γ equals the reciprocal of τ. The adopted values of τ and s were 120 msec and .16, respectively.

A typical result of this computational procedure is illustrated in Figure 6b for a tone with a rise time ϱ' (see Equation 11) of 60 msec and a duration (D) of 150 msec. Note that all our tones had decay times of 20 msec, so that the value of U_{max} was determined at the moment that the subjective envelope intersected the physical envelope during the decay portion. The duration of our tones A was 150 msec; the durations of the tones B equaled the sum of 150 msec and Δt_c (see Table 1).

The computed perceptual onsets of tones A and B for all relevant experimental combinations of rise times of tones A and B from Experiment 1 are presented in the third and fourth columns, respectively, of Table 5. The predicted effect (Δt_p) of every rise-time combination on the adjustment of the physical onset of tones B, which equals the difference between the computed perceptual onsets of tones B and A, are given in the fifth column. Inspection of the

Table 5
Computation of the Perceptual Onset by Means of a Temporal Integration Model (Schütte, 1978a) and Our Relative Threshold Model

Rise Time		Temporal Integration Model (Mean Deviation = 10.3) Perceptual Onset			Relative Threshold Model (Mean Deviation = 1.5) Perceptual Onset			Experimental Results
Tone A	Tone B	Tone A	Tone B	Δt_p	Tone A	Tone B	Δt_q	Δt_c
5	20	21	34	13	2	9	7	6
5	40	21	49	28	2	18	16	17
5	60	21	62	41	2	27	25	22
5	80	21	73	52	2	37	35	35
20	40	33	48	15	9	18	9	8
20	60	33	61	28	9	27	18	20
20	80	33	73	40	9	37	28	30
40	60	48	60	12	18	27	9	10
40	80	48	71	23	18	37	19	20
60	80	59	70	11	27	37	10	9

Note—Perceptual onsets were calculated for all tones presented in our first experiment. The parameters τ and s in the integration model were 120 msec and .16, respectively. The relative threshold level in our model was −15.4 dB. Δt_p is the experimental effect in the corresponding combinations of the rise times of tones A and B, as predicted by the temporal integration model; Δt_q is the effect as predicted by the threshold model. The observed Δt_c (see Table 1) is given in the ninth column. All values in the table are given in milliseconds.

values of Δt_p leads to the conclusion that Schütte's model is able to explain the general trend in our data from Experiment 1.

In the same vein, the perceptual onsets of tones A and B were computed in terms of our threshold model for a threshold level of −15.4 dB. The results, as well as the predicted Δt_qs, are given in Table 5.

The Predicting Power of the Two Models

We shall express the mean deviation of the observed data from the predicted data simply as the quadratic mean (rms) of the differences between the corresponding Δt_cs and Δt_ps or Δt_qs. The values of Δt_c are represented in the ninth column of Table 5. For the adopted values of τ and s of 120 msec and 16%, the geometric mean equals 10.3 msec. This mean is large, especially in view of the mean deviation of 1.5 msec that we have found between the observed data and those predicted from our relative threshold model.

However, it is possible that Schütte's model provides a better fit of our data when other values of the two parameters τ and s are chosen. It should be clear that we are interested in the power of the temporal integration model in general. It might be that the specific values of its parameters depend on such stimulus features as waveform and frequency. In our experiments, complex tones with a fundamental frequency of 400 Hz were presented, whereas Schütte used sinusoidal tones with a frequency of 2 or 3 kHz. Therefore, the mean deviations of the observed data from the predicted data were calculated for 35 combinations of τ and s, in which τ could be .1, 40, 80, 120, or 160 msec and s could be .06, .12, .14, .16, .18, .20, or .40. The mean deviations for these combinations are presented in Table 6. The smallest mean deviations are found when τ equals .1 msec and s equals about .18. From these results, it can be concluded that the predicting power of the temporal integration model, at least with regard to our data, is highest when its most important parameter, τ, is approaching 0. A temporal integration model of the perceptual onset of tones with the values of τ and s approaching 0 and .18, respectively, is, in fact, a simple threshold model (with Θ = −15 dB), as described above.

Table 6
Determination of the Optimal Values of the Parameters τ and s in the Temporal Integration Model of Schütte (1978a)

	s						
τ	.06	.12	.14	.16	.18	.20	.40
.1	8.8	3.9	3.1	1.8	1.6	2.1	11.6
40	3.5	5.8	6.9	7.8	8.9	9.7	16.6
80	4.0	7.6	8.8	9.6	10.9	11.3	15.6
120	4.7	8.4	9.4	10.2	11.0	11.6	14.2
160	5.2	8.9	9.6	10.4	11.2	12.1	13.1

Note—The mean deviations (in milliseconds) of the observed data from the predicted data are given for the presented combinations of τ and s.

Perceptual Onset and Performed Music

In studies on the temporal structure of performed music, our threshold model can be applied to determine the perceptual onsets of musical tones. When music is performed on instruments with very short rise times, like the piano, harpsichord, and drums (Gabrielsson, 1974; Povel, 1977; Sundberg & Verrillo, Note 2), the difference between the physical onset and the perceptual onset is very small. In these cases, level above threshold, too, does not have a great impact on this difference. However, when ensemble music is performed on instruments, producing tones with relatively long rise times (Rasch, 1979), such as bowed string instruments, the perceptual onset heavily depends on the relative threshold.

In such musical practice in which dynamic differences are not very large, perceptual onset is clearly affected only by the rise times and rise functions of the different instruments. This, however, is a variable with which the respective musicians can cope by adjusting their physical onset times in order to establish the appropriate timing of the perceptual onsets of their tones.

Future research should be focused on the perceptual onset of musical tones in synchronously perceived tone pairs. The sensation levels of simultaneously presented tones, especially, are dependent on the amount of auditory masking (Zwislocki, 1978). To apply our model to simultaneously produced tones, experimental results of binaural masking experiments with complex tones are needed. In addition, it would be interesting to see if our model also works in cases of complex tones consisting of partials with unequal rise times and unequal physical onsets (Freedman, 1967; Grey & Moorer, 1977) and of tones with substantially differing amplitude envelopes (Strong & Clark, 1967).

CONCLUSION

From our experimental results, we may conclude that: (1) the perceptual onsets of successively presented tones can be defined as the times at which the envelopes pass a relative threshold level; (2) within a range of 20 to 70 dB above masked or absolute threshold, the threshold for the perceptual onset lies between about 6 and 15 dB below the maximum level of the tones; and (3) a 7-dB level above masked or absolute threshold increment results in a 1-dB relative threshold decrement.

A detailed comparison of our model with a current temporal integration model revealed that, although Schütte's temporal integration model can explain the general trend of our data, it is much less powerful than our simple threshold model.

As an *explanation* of the experimental results, we propose that adaptation of the hearing mechanism to a certain relative stimulus level is responsible for perceptual onset.

REFERENCE NOTES

1. Vos, P. G. M. M. *Critical duration ratios in tone sequences for the perception of rhythmic accent and grouping* (Internal Report 79 ON 10). Nijmegen, The Netherlands: Katholieke Universiteit Nijmegen, 1979.

2. Sundberg, J., & Verrillo, V. *On the anatomy of the retard, a study of timing in music* (Quarterly Progress and Status Report 2-3). Stockholm: Speech Transmission Laboratory, 1977.

REFERENCES

Burghardt, H. Einfaches Funktionsschema zur Beschreibung der subjektiven Dauer von Schallimpulsen und Schallpausen. *Kybernetik*, 1972, **12**, 21-29.

Efron, R. Effect of stimulus duration on perceptual onset and offset latencies. *Perception & Psychophysics*, 1970, **8**, 231-234. (a)

Efron, R. The minimum duration of a perception. *Neuropsychologia*, 1970, **8**, 57-63. (b)

Efron, R. The relationship between the duration of a stimulus and the duration of a perception. *Neuropsychologia*, 1970, **8**, 37-55. (c)

Fowler, C. A. "Perceptual centers" in speech production and perception. *Perception & Psychophysics*, 1979, **25**, 375-388.

Freedman, M. D. Analysis of musical instrument tones. *Journal of the Acoustical Society of America*, 1967, **41**, 793-806.

Gabrielsson, A. Performance of rhythm patterns. *Scandinavian Journal of Psychology*, 1974, **15**, 63-72.

Grey, J. M., & Moorer, J. A. Perceptual evaluations of synthesized musical instrument tones. *Journal of the Acoustical Society of America*, 1977, **62**, 454-462.

Kirk, R. E. *Experimental design: Procedures for the behavioral sciences*. Belmont, Calif: Wadsworth, 1968.

Luce, D., & Clark, M. Durations of attack transients of nonpercussive orchestral instruments. *Journal of the Audio Engineering Society*, 1965, **13**, 194-199.

Marcus, S. M. *Perceptual centres*. Unpublished fellowship dissertation, King's College, Cambridge, 1976.

Melka, A. Messungen der Klangeinsatzdauer bei Musikinstrumenten. *Acustica*, 1970, **23**, 108-117.

Michon, J. A. *Timing in temporal tracking*. Doctoral dissertation, Institute for Perception TNO, Soesterberg, 1967.

Morton, J., Marcus, S., & Frankish, C. Perceptual centers. *Psychological Review*, 1976, **83**, 405-408.

Munson, W. A. The growth of auditory sensation. *Journal of the Acoustical Society of America*, 1947, **19**, 584-591.

Plomp, R. Hearing threshold for periodic tone pulses. *Journal of the Acoustical Society of America*, 1961, **33**, 1561-1569.

Plomp, R., & Bouman, M. A. Relation between hearing threshold and duration for tone pulses. *Journal of the Acoustical Society of America*, 1959, **31**, 749-758.

Povel, D. Temporal structure of performed music, some preliminary observations. *Acta Psychologica*, 1977, **41**, 309-320.

Rasch, R. A. Synchronization in performed ensemble music. *Acustica*, 1979, **43**, 121-131.

Schütte, H. *Bestimmung der subjektiven Ereigniszeitpunkte aufeinanderfolgender Schallimpulse durch Psychoakustische Messungen*. Doctoral dissertation, Technical University, Munich, 1977.

Schütte, H. Ein Funktionsschema für die Wahrnehmung eines gleichmässigen Rhythmus in Schallimpulsfolgen. *Biological Cybernetics*, 1978, **29**, 49-55. (a)

Schütte, H. Subjektiv gleichmässiger Rhythmus: Ein Beitrag zur zeitlichen Wahrnehmung von Schallereignissen. *Acustica*, 1978, **14**, 197-206. (b)

Stott, L. H. Time order errors in the discrimination of short tonal durations. *Journal of Experimental Psychology*, 1935, **18**, 741-766.

Strong, W., & Clark, M. Synthesis of wind-instrument tones. *Journal of the Acoustical Society of America*, 1967, **41**, 39-52.

Vos, P. G. M. M. *Waarneming van Metrische Toonreeksen*. Doctoral dissertation, University of Nijmegen, The Netherlands, 1973.

Woodrow, H. The temporal indifference interval determined by the method of mean error. *Journal of Experimental Psychology*, 1934, **17**, 167-188.

Woodrow, H. The effect of practice upon time-order errors in the comparison of temporal intervals. *Psychological Review*, 1935, **42**, 127-152.

Wundt, W. *Grundzüge der physiologischen Psychologie* (Vol. 3, 5th ed.). Leipzig: W. Bertelmann, 1903.

Zwicker, E. Subjektive und objektive Dauer von Schallimpulsen und Schallpausen. *Acustica*, 1970, **22**, 214-218.

Zwislocki, J. Theory of temporal auditory summation. *Journal of the Acoustical Society of America*, 1960, **32**, 1046-1060.

Zwislocki, J. J. Temporal summation of loudness: An analysis. *Journal of the Acoustical Society of America*, 1969, **46**, 431-441.

Zwislocki, J. J. Masking: Experimental and theoretical aspects of simultaneous, forward, backward, and central masking. In E. Carterette & M. Friedman (Eds.), *Handbook of perception* (Vol. 4, chap. 8). New York: Academic Press, 1978.

(Received for publication July 17, 1980;
revision accepted January 5, 1981.)