Detection of pitch motion in complex sounds

Versfeld, N.J.

Award date:
1988

DETECTION OF PITCH MOTION IN COMPLEX SOUNDS

N.J. Versfeld

Instituut voor Perceptie Onderzoek
Postbus 513, 5600 MB EINDHOVEN

NV/nv
02.03.1988

Rapport no. 636

Detection of pitch motion in complex sounds

N.J. Versfeld

Eindhoven, March 1988
Coach: Dr. A.J.M. Houtsma
This report represents partial fulfillment of the requirements for the degree of Master of Science in Physics

Abstract

Perception of pitch motion in complex sounds was studied in three experiments. The first experiment dealt with the perception of global pitch movement in random-chord sequences. The subject's behaviour in identifying overall pitch motion can be accounted for by a model in which the ultimate pitch percept is determined by a function of elementary events in the sequence. The second experiment dealt with the perception of an up- or downward musical scale against otherwise randomly changing tones. It appears that the subject adopts the strategy of tracking only one of two possible scales, from which he can deduce which scale actually was present. The third experiment dealt with the detectability of a frequency jump in a flat 20-component spectrum, due to changes in amplitude in adjacent components of successive time frames. It appears that the 70.7%-correct threshold varies from 1.2 to 1.9 dB when the subject does not know a priori where the change in amplitude occurs, and decreases even to 0.25 dB in some conditions when he does know the place.

Contents

1 Introduction

The problem of pitch perception for sinusoidal tones has received considerable attention in the literature on psychoacoustics. Relations have been established between subjective pitch and objective acoustic variables such as the tone's frequency (Stevens and Volkmann, 1940), its intensity (Stevens, 1935; Verschuure and van Meeteren, 1975), its duration (Doughty and Garner, 1948), its temporal envelope (Hartmann, 1978; Rossing and Houtsma, 1986) and the presence and strength of other interfering sounds (Terhardt and Fastl, 1971; Larkin, 1978).

Pitch perception for complex tones has received much attention as well during the past few decades. The pitch of a harmonic complex tone is not merely determined by its fundamental frequency, but is to a very large extent influenced by its harmonics. Tones of church bells, orchestral chimes or filtered tones from common musical instruments very often produce a clear, unambiguous pitch sensation without the presence of any acoustic energy at the fundamental frequency. It has been established over the years that this so-called *missing fundamental* percept is not accounted for by difference-tone distortion or by periodicity detection of interference patterns in the ear, but is the result of the way in which our brain processes the neural transformations of sounds from our two ears. Detailed reviews of this research have been given by de Boer (1976) and Scharf and Houtsma (1986). When two complex tones, each comprising a few harmonics, sound simultaneously, the pitches corresponding to each tone can usually be heard as well (Beerends and Houtsma, 1986). The exact perceptual limit to the number of simultaneous tones or pitches that can be correctly perceived is

not known, and is probably very dependent on training (Doehring, 1971).

Perception of pitch sequences is another problem that has received a fair amount of attention in the past. A temporal sequence of two tones, called a *melodic interval*, is typically perceived in a categorical manner (Burns and Ward, 1978). Eiting (1984) showed that recognition of three-note sequences occurs on the basis of contour (e.g., up-down-up) as well as on the size of successively perceived melodic intervals. Deutsch (1980) studied listeners' retention capacity for longer tonal sequences and found that it depended greatly on the degree of hierarchical structure in the sequence. For very fast tonal sequences (more that 15 notes per second) a stream of notes may become perceptually separated into two or more parallel streams, each forming a melody (Bregman and Campbell, 1971; Dowling, 1973; van Noorden, 1975).

Less attention has been paid to the study of sequences of simultaneous tones. One of the few situations that has received attention is the dichotic conflict situation in which two different melodies are simultaneously presented to a subject, one to each ear (Kimura, 1964; Deutsch, 1975; Butler, 1979). The limits of our auditory system to perceive a melody against a background of other potentially interfering tones without dichotic separation remains largely unknown. This is somewhat unfortunate because it is just this situation that is encountered most often when we listen to music.

This report deals with the study of pitch motion perception in sequences of simultaneous tones. In experiments with such complex changing sounds grouping mechanisms like frequency proximity, timbre, amplitude or good continuation play an important role in identification processes (Deutsch,

1982, section IV). In order to allow quantitative measurements and data analyses, a review of signal detection theory as applied to forced choice fixed-stimulus and adaptive procedures will be given (Chapter 2).

The first experiment (Chapter 3) deals with the perception of randomly changing elements, moving according to a probability scheme, which evokes a global pitch movement. In this type of experiment the issue of correspondence, i.e., the question of which elements in one time frame correspond to which elements in the next frame(s), plays the most important role. The aim of this experiment is to discover if the results can be explained with a theory that uses only the shortest jumps in frequency and time (*local* correspondence, Bell & Lappin, 1973; Allik & Dzhafarov, 1984).

The second experiment (Chapter 4) deals with the perception of a discretely moving signal against a background of other randomly changing elements. This type of experiment differs from the preceeding one since now a signal and noise (i.e. the randomly moving elements) are to be separated in the identification process in order to perceive the motion direction of the signal, whereas in the previous experiment this motion percept is part of the whole. Therefore not only correspondence but good continuation are important in the identification process as well. The aim of this experiment is to reveal some elements of the detection process of signals in a noise-like background.

The third experiment (Chapter 5) deals with the perception of a motion due to a change in amplitude of two successive components in a complex spectrum. The aim of this experiment was to determine the threshold of the change as a function of the place of the component in the complex and

as a function of the shape of the complex. In particular we will investigate whether small changes in details of a complex sound are detected as changing components or as changes in overall pattern.

2 Theory

2.1 Signal Detection Theory in 2-Alternative Forced Choice Procedures

Signal detection theory, developed largely for radar applications during the Second World War, has proven very useful for the description of auditory processes as well (Green & Swets, 1966). Auditory discrimination of two very similar sounds, absolute identification of several sounds, detection of a sound signal against a background of noise, and even magnitude estimation and category scaling can all be described and parametrized by this theory. Since in this study we will deal with stimuli in which either an up- or downward frequency motion has to be detected, signal detection theory will play a central role in this report and will therefore be reviewed in this chapter.

In a typical 2-alternative forced-choice auditory test, a subject is presented on each trial one out of two possible stimuli, and asked to give one out of two possible responses. The stimulus could be a tone buried in noise versus noise without a tone, and the response "tone present" versus "tone absent". In this study the general situation will be stimuli with an up- or downward frequency motion, and responses of "pitch up" or "pitch down". The possible stimulus-response combinations, together with their short-hand notation, have been listed in Table 2.1.1.

Stimulus	Subject's response	Number of occurrence
up	*up*	N_{uu}
up	*down*	N_{ud}
down	*up*	N_{du}
down	*down*	N_{dd}

Table 2.1.1.: Results of a 2-alternative forced choice experiment.

A perception and decision process is modeled in a black-box fashion. We assume that there exists a real random decision variable X $(-\infty < X < \infty)$ with the property that each stimulus, when presented to a subject, gives rise to particular value X_s. Futhermore there exists some internal criterion X_0 at the X-axis, so that if $X_s < X_0$ the subject responds *down* and if $X_s \geq X_0$ the subject responds *up*. We assume that X_0 is constant during the experiments.

Due to all kinds of human factors (degree of concentration ect.), physical noise in the stimulus (e.g. the presence of background noise) or physiological noise (neural firing), the same stimulus does not always give rise to the same value of X_s. Due to the discrete nature and the huge amount of neural firing we assume, on grounds of the Central Limit Theory, that the distribution of X_s over the X-axis for one type of stimulus is Gaussian. The mean of the Gaussian curve represents the expected value for X_s, the variance represents the associated amount of noise. Since there is no reason to believe that the amount of noise in the response of an *up*-stimulus differs from the amount of noise of a *down*-stimulus, we assume that the variance

in the response of the two types of stimuli are equal thus $\sigma_u = \sigma_d = \sigma$. If we denote the mean of the *up*-stimuli with X_u and the mean of the *down*-stimuli with X_d, we get four expressions for N_{uu}, N_{ud}, N_{du} and N_{dd} (see figure 2.1.1):

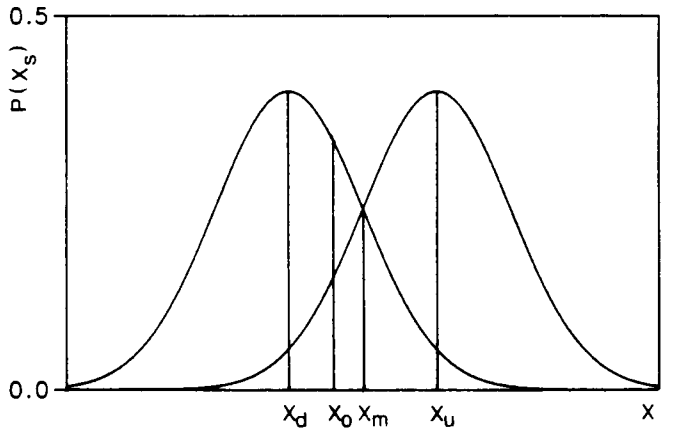


Figure 2.1.1: Distribution of X_s over the X-axis.

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{X_0} e^{-\frac{1}{2}(\frac{X-X_d}{\sigma})^2} dX = \frac{N_{dd}}{N_{dd} + N_{du}} =: P_{dd}, \tag{1}$$

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{X_0}^{\infty} e^{-\frac{1}{2}(\frac{X-X_d}{\sigma})^2} dX = \frac{N_{du}}{N_{dd} + N_{du}} =: P_{du}, \tag{2}$$

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{X_0} e^{-\frac{1}{2}(\frac{X-X_u}{\sigma})^2} dX = \frac{N_{ud}}{N_{ud} + N_{uu}} =: P_{ud} \text{ and} \tag{3}$$

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{X_0}^{\infty} e^{-\frac{1}{2}(\frac{X-X_u}{\sigma})^2} dX = \frac{N_{uu}}{N_{ud} + N_{uu}} =: P_{uu}, \tag{4}$$

where e.g. P_{ud} is the chance that the subject responds *down* to an *up*-stimulus. Since $P_{uu} + P_{ud} = 1$ and $P_{du} + P_{dd} = 1$, we actually deal with only two independent equations. For convenience we choose equations (1) and (3). We define the sensitivity d' as the normalized distance between X_u and X_d:

$$d' := \frac{X_u - X_d}{\sigma}. \tag{5}$$

d' indicates to what degree a subject tends to confuse both signals. If d' is large both signals are practically not confused; if $d' = 0$ the signals can't be distinguished and performance is at chance-level. In equation (1) and (3) we make the transformation

$$t = \frac{X - X_u}{\sigma}, \tag{6}$$

and substitute d', so we get

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-d} e^{-\frac{1}{2}(t+d')^2} dt = \mathrm{P_{dd}}, \tag{7}$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-d} e^{-\frac{1}{2}t^2} dt = \mathrm{P_{ud}}, \tag{8}$$

where we defined d as the normalized distance between X_u and X_0:

$$d := \frac{X_u - X_0}{\sigma}. \tag{9}$$

The Normal Probability Function is defined as (Abramowitz & Stegun, 1956, sect. 26.2):

$$\Phi(z) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}u^2} du. \tag{10}$$

In order to write equation (7) as such a function we make the transformation $v = t + d'$ and get:

$$\Phi(d' - d) = \mathrm{P_{dd}} \tag{11}$$

$$\Phi(-d) = \mathrm{P_{ud}}. \tag{12}$$

Or, since d and d' are unknown:

$$d = -\Phi^{-1}(\mathrm{P_{ud}}), \tag{13}$$

$$d' = \Phi^{-1}(P_{dd}) - \Phi^{-1}(P_{ud}), \tag{14}$$

where $\Phi^{-1}(z)$ is the inverse of the Normal Probability Function. We now have found an expression for d and d'.

Suppose we have an ideal subject. This subject tries to maximize his fraction of correct responses P_c:

$$P_c := \frac{N_{uu} + N_{dd}}{N_{uu} + N_{ud} + N_{du} + N_{dd}}. \tag{15}$$

Thus he tries to find an X_m which satisfies:

$$\frac{\partial P_c}{\partial X_0}|_{X_0 = X_m} = 0. \tag{16}$$

With the aid of equations (1) and (4) we write:

$$P_c = \frac{(N_{du} + N_{dd})\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{X_0} e^{-\frac{1}{2}(\frac{X - X_d}{\sigma})^2} dX}{N_{uu} + N_{ud} + N_{du} + N_{dd}}$$

$$+ \frac{(N_{ud} + N_{uu})\frac{1}{\sigma\sqrt{2\pi}} \int_{X_0}^{\infty} e^{-\frac{1}{2}(\frac{X - X_u}{\sigma})^2} dX}{N_{uu} + N_{ud} + N_{du} + N_{dd}}, \tag{17}$$

and

$$\frac{\partial P_c}{\partial X_0}|_{X_0 = X_m} = \frac{(N_{du} + N_{dd})\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{X_m - X_d}{\sigma})^2}}{N_{uu} + N_{ud} + N_{du} + N_{dd}}$$

$$- \frac{(N_{ud} + N_{uu})\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{X_m - X_u}{\sigma})^2}}{N_{uu} + N_{ud} + N_{du} + N_{dd}} = 0. \tag{18}$$

Thus

$$(N_{ud} + N_{uu})\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{X_m - X_u}{\sigma})^2} = (N_{du} + N_{dd})\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{X_m - X_d}{\sigma})^2} \tag{19}$$

Solving equation (19) for X_m gives us

$$X_m = \frac{1}{2}(X_d + X_u) + \frac{\sigma^2}{X_u - X_d} \ln(\frac{N_{uu} + N_{ud}}{N_{du} + N_{dd}}), \tag{20}$$

or, if we normalize X_m to σ and express equation (20) in terms of d':

$$\frac{X_m}{\sigma} = \frac{X_d + X_u}{2\sigma} + \frac{1}{d'} \ln(\frac{N_{uu} + N_{ud}}{N_{du} + N_{dd}}). \tag{21}$$

In the simple case, where the number of *up*-stimuli equals the number of *down*-stimuli, we see that X_m is the mean of X_u and X_d and lies at the point of intersection of the two Gaussian curves (see figure 2.1.1).

However, a non-ideal subject uses the criterion X_0 instead of X_m. The difference between these criteria is a measure for the subject's tendency to answer more frequenty, say, *up*. This tendency is called the *bias* β, and is defined as (Lippman *et al.*, 1976):

$$\begin{aligned} \beta &:= \frac{X_0 - X_m}{\sigma} \\ &= \frac{1}{2}d' - d - \frac{1}{d'} \ln(\frac{N_{uu} + N_{ud}}{N_{du} + N_{dd}}). \end{aligned} \tag{22}$$

If there exists no preference, $\beta = 0$; if $\beta > 0$, there is a tendency for *down*-responses, if $\beta < 0$, a tendency for *up*-responses.

The definition of d' (being the distance between X_d and X_u) is a more appropriate quantity for the subject's ability to distinguish two signals than P_c, since d' is freed from bias. So it doesn't matter if there is some bias, for we can eliminate this from the data.

Mention that up to now we dealt with only two stimuli (*up* and *down*), differing only in pitch motion direction. The difference variable Δ of the signal was not varied. One can imagine that, when varying Δ, d' and β may vary as well.

2.2 Method Of Fixed Levels

In the Method of Fixed Levels, 2-alternative Forced Choice experiments are performed at only one value of the difference variable Δ, chosen beforehand by the experimenter. This results in one point in the d'-Δ-plane. By taking some more difference values levels more points are obtained and one gets an impression of d' as a function of Δ.

Many experiments (Green & Swets, 1966) have shown that often d' is a linear function of Δ. If indeed this is the case, we only need two values of d' by different values of Δ and we directly know the performance for all values of Δ. If we want to fit the line $l : d' = a + b\Delta$ we have to find the slope b and the intersection point with the d'-axis a. For covenience, we define $r := -a/b$ as the point of intersection of l with the Δ-axis. The slope b is a measure for the detectability or amount of noise in the human system or in the signal.If b is small, only at high signal intensities the signal is easy perceived; if b is large, detectability already becomes easy at low intensities. The offset r is a measure for the parameter uncertainty in the human system (Tanner *et al.*, 1970, p.88 and Jeffress, 1970, p.109). If we deal with an ideal observer r should be zero, but a human observer sometimes causes r to be positive. This means that performance stays at chance level upto $\Delta = r$ (i.e. $d' = 0$ for $0 \leq \Delta \leq r$). From there on d' increases linearly with Δ, following l.

2.3 Adaptive Procedures

If we assume that d' is a linear function of Δ and the offset level $r = 0$, we only need one point at the d'-Δ-plane to estimate the line l. A very quick

and effective way to estimate one point is the adaptive procedure (Levitt & Rabiner, 1967; Levitt, 1971 or Shelton, 1982). In Chapter 5 we will use the transformed up-down method with two steps, as described in Levitt (1971).

This procedure starts with the presentation of two stimuli with a high probability of a correct response (or, in other words, two easy stimuli with large difference Δ). If both the subject's responses are correct Δ is decreased one step, and again two stimuli are presented. If again both responses are correct, Δ is decreased and this procedure is continued until one or two negative responses are obtained. Instead of a decrease, we now increase Δ and each time an incorrect answer results in an increase for Δ. If again two responses are correct Δ will be decreased, et cetera. A typical data record is shown in figure 2.3.1. The moment a decrease in Δ is followed by an increase in Δ or vice versa is called a *reversal.* The amount of in- or decrease of Δ is called a *step.* The values of Δ tend to converge to that value for Δ_0 at which the probability of two correct responses equals the probability of one or two incorrect responses. If we denote $P(\Delta)$ as the probability of a correct response to a stimulus with value Δ, then

$$\begin{aligned} P(\text{two correct responses}) &= \frac{1}{4}P(\Delta_0)^2 \\ P(\text{one correct and one incorrect response}) &= \frac{2}{4}P(\Delta_0)\cdot(1-P(\Delta_0)) \\ P(\text{two incorrect responses}) &= \frac{1}{4}(1-P(\Delta_0))^2 \end{aligned}$$

Thus

$$\frac{1}{4}P(\Delta_0)^2 = \frac{2}{4}P(\Delta_0)\cdot(1-P(\Delta_0)) + \frac{1}{4}(1-P(\Delta_0))^2 \qquad (23)$$

This gives $P(\Delta_0) = \frac{1}{2}\sqrt{2} \approx 0.707$. If we average over all (except the first

four) reversals we have a good approximation of the difference Δ at which the chance at a correct response is 70.7 % . This point corresponds with $d' = 1.0889$ if $r = 0$.

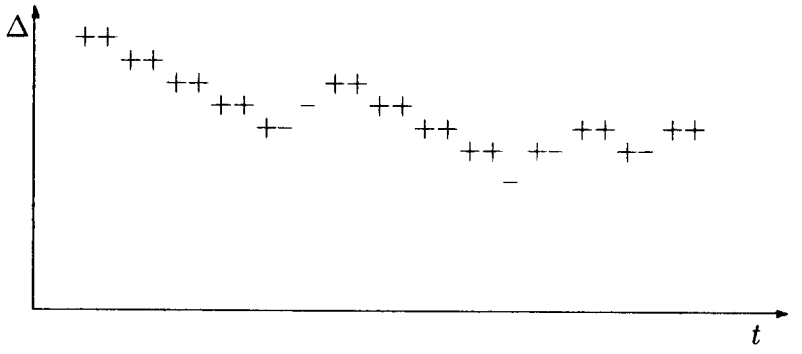


Figure 2.3.1: Typical data record of a transformed up-down method with two steps. "+" and "–" denote a correct and incorrect response respectively. This figure shows five reversals.

3 Pitch Motion with Random Chord Sequences

3.1 Introduction

This chapter deals with the perception of global pitch motion in so-called Random Chord Sequences (RCSs). In hearing little attention has been paid to the study of sequences of simultaneous sounding tones.

In research on vision, however, the problem of motion detection has been given a fair more amount of attention. When (in vision) several dots are presented repeatedly, but each time in another configuration, some kind of movement might appear. In experiments with these moving random-dot patterns or *cinematograms* (Julesz, 1971) the issue of correspondence, i.e. the question of which elements in one time frame correspond (perceptively) to which elements in the next frames, becomes quite complicated. Much visual behaviour, though, can be accounted for with a so-called *local* correspondence model in which perceptual correspondence is limited to single spatial jumps in succesive time frames only (Bell & Lappin, 1973).

A special case of a random-dot pattern is the *Circular Random Cinematogram* (CRC) first applied by Allik & Dzhafarov (1984). Their CRC consisted of twelve light elements grouped circularly at the 5-minute marks on the face of a clock, where each light elements could be either on or off. A sequence of random circular displays often evokes an apparent clockwise or counterclockwise rotation percept, reflecting the so-called *phi-phenomenon* (Wertheimer, 1912; Korte, 1915), or *reversed phi-phenomenon* (Anstis, 1970). Perceptual identification data could well be accounted for

with a strictly local and short-range model in which only jumps between successive display elements in successive time frames play a role.

The study presented in this chapter is an acoustic analog of Allik and Dzhafarov's (circular) random cinematogram experiment. It deals with the perception of apparent global pitch movement in a so-called *random-chord sequence* (RCS), the acoustic equivalent of a CRC. An RCS is a sequence of pure- or complex-tone clusters in which, under certain conditions, a global upward or downward pitch movement can be heard. The topic of this study is to find out if in hearing global pitch movement as in vision can be explained with a local correspondence model, and if so, to what degree subjects are able to identify this pitch motion.

One part of the experiments described in this chapter was performed by Jüri Allik and Joan Ross in Tallinn, E.S.S.R; the other part by the author. The theory descibed in this chapter was developed by Jüri Allik and Ethibar Dzhafarov (U.S.A.) for the visual experiments and extended for hearing by Ethibar Dzhafarov, Aad Houtsma and the author. All calculations were performed by the author. This chapter is written by Aad Houtsma and the author and is to be published in the *Journal of Experimental Psychology: Human Perception and Performance.*

3.2 Random-Chord Sequences

A single complex tone is characterized by its fundamental frequency, its spectral envelope and phase function, and its duration. One such tone is called an *element.* An element can be in one of two possible states: sounding (*on*) or silent (*off*). Several such elements grouped together, each with a different fundamental frequency, form a *chord.* Since each tone has only two states, M different tones can form 2^M differently-sounding chords. A random temporal sequence of two or more of these 2^M chords is called a *Random-Chord Sequence* or RCS.

The states of elements in an RCS are determined in the following manner. First, a direction of frequency motion is chosen. This can be either upward (N=1) or downward (N=-1). Next, the states of the M elements of the first chord (or *frame*) are determined randomly, resulting in an average of 50% of the first-frame elements being in the *on* state. The states of elements in the second frame depend on the states in the first frame. If N=1, each element (*on* or *off*) of the first chord is connected with the element of the second chord that is one frequency step higher. If N=-1, it is connected with the second-chord element one frequency step lower. The connection implies that the state of each element in the second chord will follow the state of the element in the first chord with which it is connected with a probability P which is called the *State Repetition Probability* or SRP. If, for instance, the states of the lowest and highest-but-one tones of the first chord are *on* and N=1 and SRP=1, the next-lowest and the highest notes of the second chord will be *on* as well. If SRP=0, however, the next-to-lowest and highest notes of the second chord would be *off* in this example.

During an RCS the direction of motion N and the SRP are kept constant. States of elements in the third chord are determined from element states in the second chord, etc. To obtain a circular scheme similar to the visual displays used by Allik and Dzhafarov (1984), extreme chord elements are also connected: for N=1 the highest note of the ith frame is connected with the lowest note of the $i+1$th frame, and for N=-1 the lowest note of the ith frame with the highest note of the $i+1$th frame. Three examples of RCSs are illustrated in Figure 3.2.1.

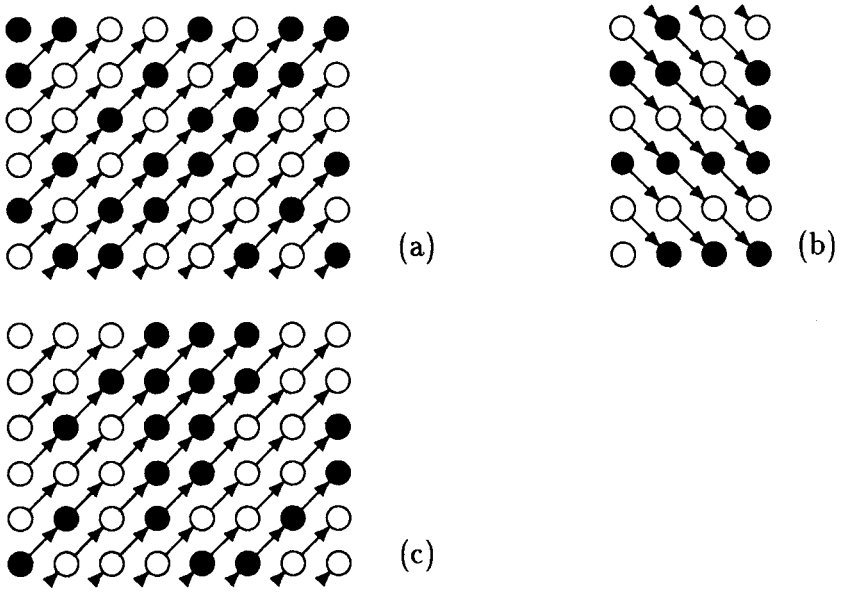


Figure 3.2.1: Three examples of random-chord sequences. (a) N=1, SRP=1, circular scheme, 6 elements, 8 frames; (b) N=-1, SRP=0, circular scheme, 6 elements,4 frames; (c) N=1, SRP=0.8, circular scheme, 6 elements, 8 frames. An open circle denotes an off-element; a closed circle denotes an on-element.

3.3 Experiments

Two series of experiments are reported. They both represent attempts to measure subjects' ability to identify the direction of perceived global pitch motion for RCSs of two, four, five and eight sequential chords. The first series of experiments (Experiment A) investigated the effect of the sound used to represent chord elements. This series was carried out with four subjects at the Institute for Perception Research in The Netherlands, using a P 857 mini-computer and a 12-bit D/A converter to compute, store, synthesize and present stimulus sequences. In the second series of experiments (Experiment B) the effect of chord composition was investigated. These experiments were carried out at the Institute of Language and Literature of the Estonian Academy of Sciences in Tallinn, E.S.S.R., with two subjects, using an EC 1010 mini-computer with 12-bit D/A converter. Both sets of experiments are presented separately.

3.3.1 Stimuli

In Experiment A there were always six elements (tones) in a chord, each of which could be either on or off. Tones were either pure sinusoids, sawtooth waves or so-called *Shepard* tones. Of the sawtooth waves, only the first four harmonics were included. Shepard tones (Shepard, 1964) are complex tones with octave harmonics and a fixed bell-shaped spectral weight function. They have the property that transposition by one or more octaves always yields the same physical tone, and therefore the same perceived pitch. This circular pitch property was used to make RCSs perceptually completely circular. Wave samples and spectra of the three sounds are

shown in Figure 3.3.1. Sequences of two, four and eight 6-element chords were used. For sinusoidal and sawtooth-wave sounds the fundamental frequencies of elements were chosen at 370, 392, 415, 440, 466 and 494 Hz, representing an inter-tone spacing of one equally-tempered semitone. Shepard tones were tuned to 262, 294, 330, 370, 415 and 466 Hz, two semitones apart.

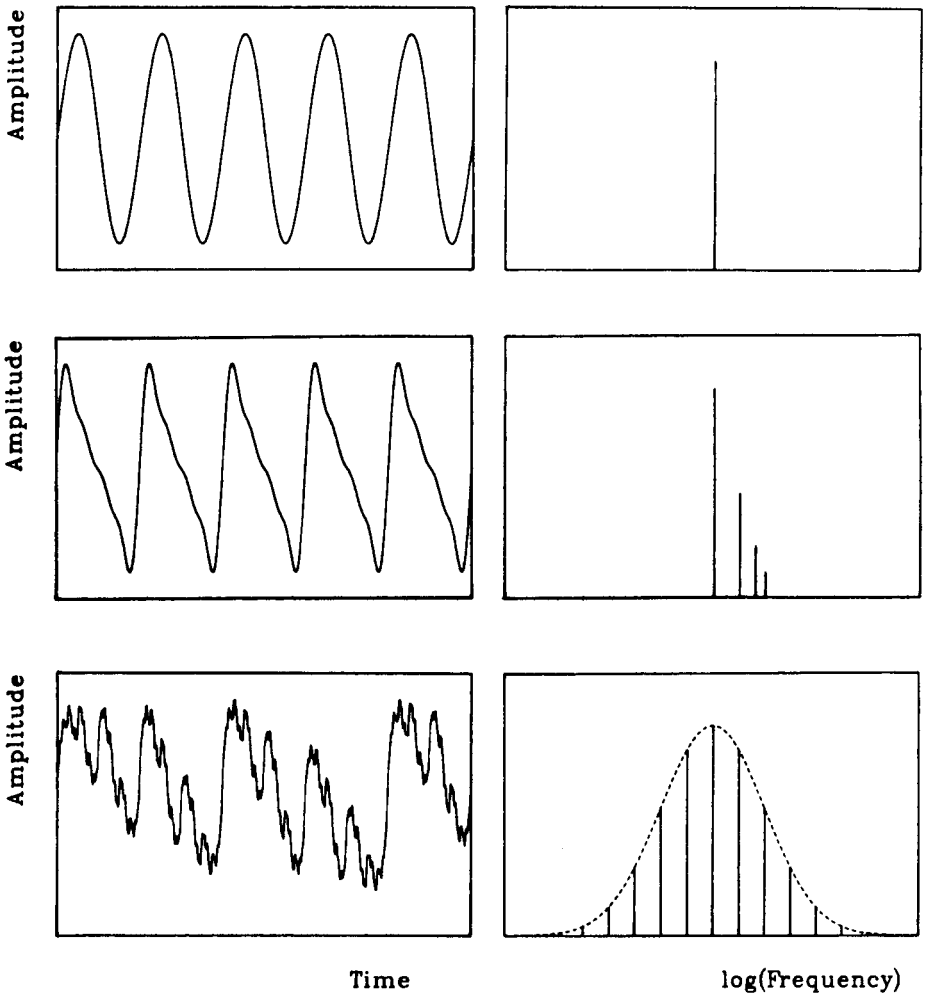


Figure 3.3.1: Samples of waveforms and spectra for the three types of sounds used in Experiments A and B. (a) Sinusoid, (b) Sawtooth-like wave, (c) Shepard tone.

The duration of each chord was 250 ms which included a 20-ms on and a 20-ms off ramp. There were no inter-chord silent periods. The SRP varied between trials from 0.0 to 1.0 in steps of 0.1.

In Experiment B there were either six or eight elements (tones) in a chord. The tones were always sawtooth-like complex tones comprising 4 harmonics, as shown in Figure 3.3.1(b). In one experiment six chord elements were arranged in semitone increments with fundamental frequencies at 370, 392, 415, 440, 466 and 494 Hz, as in Experiment A. In another experiment fundamental frequencies were arranged to form a dominant-seventh chord of frequencies 196, 247, 294, 349, 392 and 494 Hz. In a third and fourth experiment, in which eight elements per chord were used, fundamentals were arranged in quarter tones (440, 453, 466, 480, 494, 508, 523, 539 Hz) and as a dominant-seventh chord (196, 247, 294, 349, 392, 494, 587, 696 Hz). All chords had a duration of 250 ms which included a 20-ms on and 20-ms off ramp. In the first two experiments RCSs of two, four and eight chords were used; in the last two experiments only RCSs of five chords were used.

3.3.2 Method

In Experiment A all sessions started with a determination of the subject's hearing threshold for the chords to be used. The subject, who was seated in a double-walled sound-insulated chamber and received the stimuli binaurally through headphones, adjusted the intensity of an intermittent 440-Hz sinusoidal, sawtooth or Shepard tone to detection threshold. All stimuli in the experiment were presented 20 dB above this empirically estab-

lished level. This rather low level was chosen to avoid as much as possible confounding effects of aural combination tones (Zwicker, 1955; Goldstein, 1967). After presentation of each RCS the subject indicated whether the perceived global pitch motion was upward or downward by pressing one of two buttons on a response box. There was no reponse time limit and no feedback was provided. From each subject 100 trials were collected for each sound condition (i.e., 2-, 4-, and 8-frame stimuli combined with each of the three waveforms) and at each SRP value. Feedback was not provided because there were no right or wrong answers. The subject's task was to indicate the subjectively perceived direction of pitch motion and not the physical direction of frequency motion N.

In Experiment B stimuli were synthesized digitally, played through a 12-bit D/A converter and stored on magnetic tape. Subjects were seated in a quiet room and received the stimuli through loudspeakers. They were allowed to choose a comfortable sound level. After presentation of each RCS the subject indicated whether the perceived global pitch motion was upward or downward by writing the letters Y or A on a score sheet. The response time allowed was 2 s. There were 100 trials in the first and second, and 200 trials in the third and fourth experiment for each sound condition and each SRP value. No feedback was provided.

3.3.3 Subjects

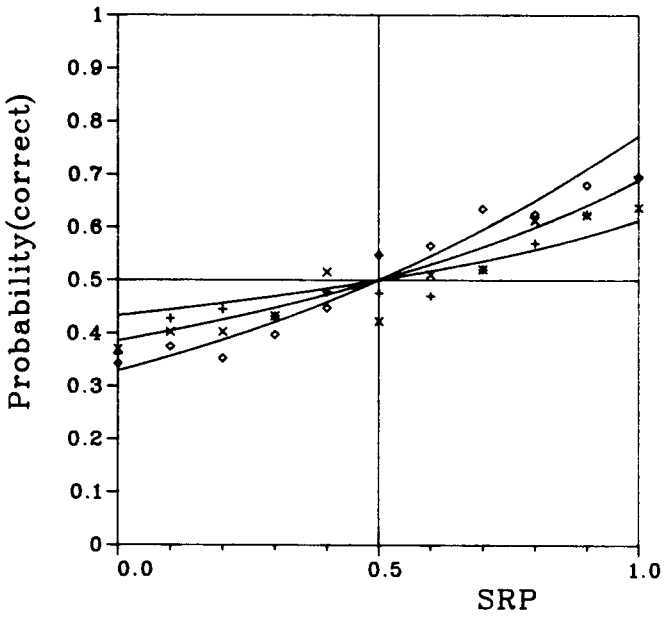
In Experiment A all RCSs with sinusoidal and sawtooth wave tones were judged by four subjects AH, BE, JJ and NV all of whom had some musical training and experience. RCSs with Shepard tones were judged by three

subjects AH, JJ and NV. All subjects had normal hearing.

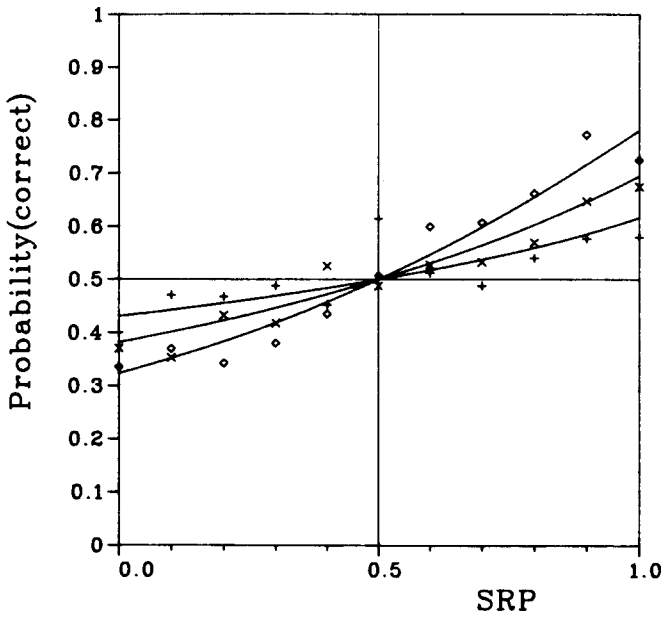
In Experiment B all RCSs were judged by two subjects JR and MR, who both had some musical training and had normal hearing.

3.4 Results

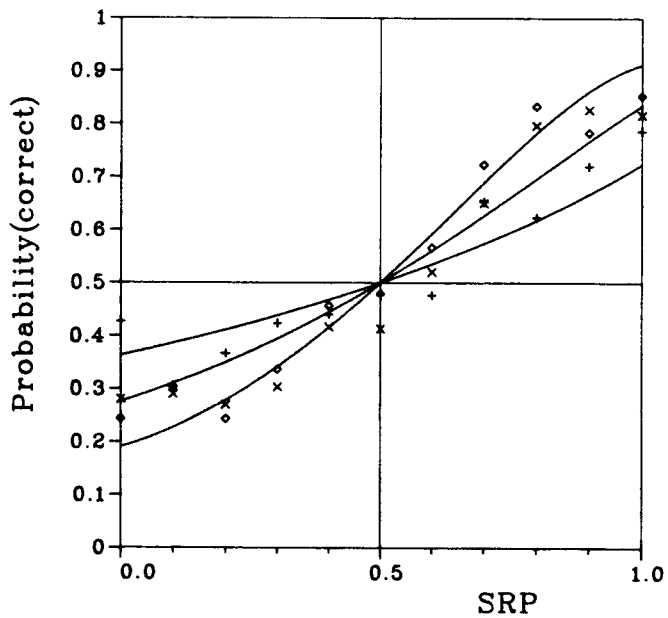
The results of Experiment A are summarized by the data points shown in Figure 3.4.1. For each of the sounds used to represent chord elements, the fraction Pc of trials is shown at each SRP value for which the direction of perceived global pitch motion agreed with the chosen direction of frequency motion N. Because all four subjects appeared to behave similarly, their data were pooled. Each data point, in which a different symbol designates measurements with two-chord, four-chord and eight-chord sequences, therefore represents $4 \times 100 = 400$ experimental trials. The psychometric functions fitted through these points are model predictions and will be discussed in the next section.



(a)



(b)



(c)

Figure 3.4.1: Experimental and theoretical results of Experiment A. (a) Results obtained with sinusoidal tones, $V_{nj} = 14.6$*; (b) results with sawtooth waves,* $V_{nj} = 13.6$*; (c) results with Shepard tones,* $V_{nj} = 4.0$*. +: 2-chord, ×: 4-chord, ◇: 8-chord sequence.*

The results of Experiment B are shown comprehensively in Figure 3.4.2. The first panel (a) shows results obtained from one subject with sequences of two, four and eight chords comprising up to six sawtooth wave tones arranged in semitone steps. This condition was comparable to that of Experiment A, part 2, the results of which were shown in Figure 3.4.1(b). The second panel of Figure 3.4.2, (b), shows the results from the same subject for an arrangement of chord elements according to a dominant-seventh chord. Figure 3.4.2(c) shows the averaged results of two subjects obtained with five-chord sequences of eight elements arranged either in quarter-tone steps or in steps that form a dominant-seventh chord. Data points in Figures 3.4.2(a) and (b) represent 100 trials and those in Figure 3.4.2(c) 200 trials each. The functions shown represent model fits to be discussed later.

Observation of the data of Figures 3.4.1 and 3.4.2 reveals the following general tendencies:

(a) Although the functions formed by the data points appear to pass through the point ($Pc = 0.5$, SRP=0.5), as is expected, the functions show a distinct lack of odd symmetry about this point. For SRP values smaller than 0.5, Pc is generally closer to 0.5 than for SRP values larger than 0.5. All functions appear to satisfy the inequality:

$$Pc(\mathrm{SRP}) + Pc(1 - \mathrm{SRP}) > 1 \qquad (1)$$

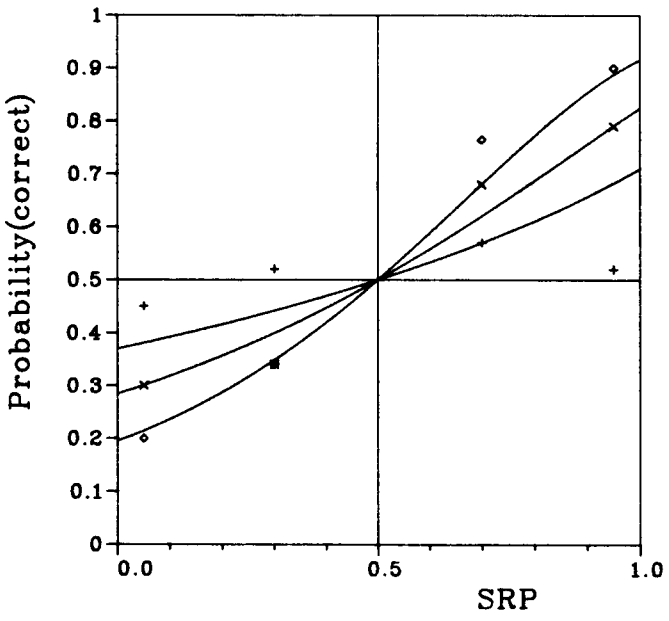
The probability that this asymmetry is accidental was calculated to be smaller than 0.06% at the 95% confidence level for the data of Experiment A.

(b) For SRP values smaller than 0.5, subjects always perceive a global pitch

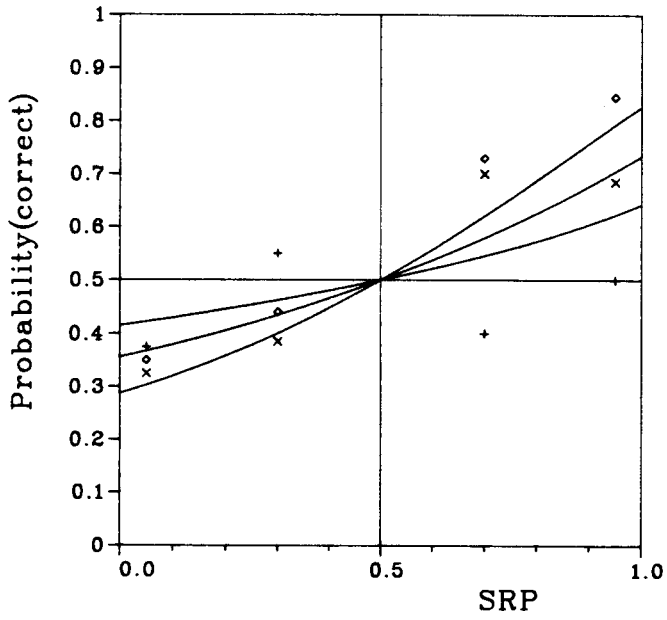
movement opposite to the actual direction of frequency movement. This apparent "pitch direction reversal" phenomenon holds for all chord and signal conditions. It is analogous to the "direction reversal" phenomenon reported by Allik and Dzhafarov (1984) and the "reversed phi" phenomenon reported by Anstis (1970).

(c) Both positive and the negative portions of the psychometric functions take on more extreme values when the number of chords in a RCS is increased or the sound used to represent chord elements is spectrally enriched.

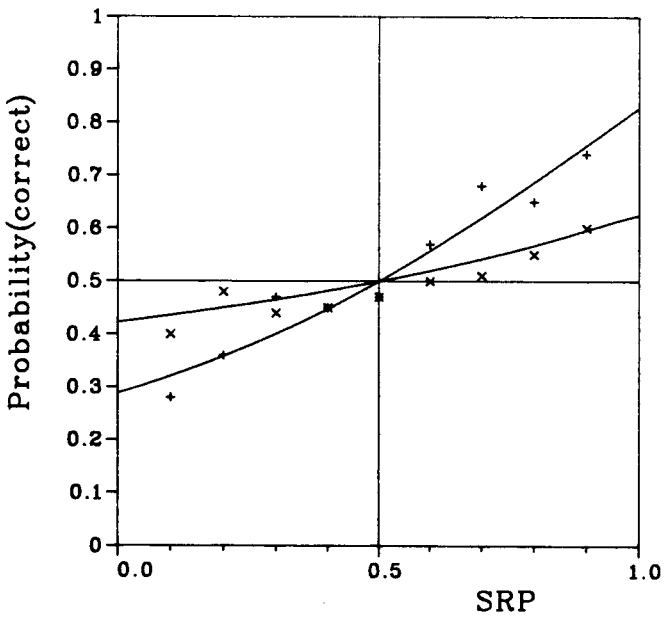
(d) The functions obtained reach more extreme values of Pc when chord elements are regularly spaced (whole tones, semitones, quarter tones) than when they are irregularly spaced, as is the case with the dominant-seventh chord arrangement.



(a)



(b)



(c)

Figure 3.4.2: Experimental and theoretical results of Experiment B performed with sawtooth waves. (a) Element spacing in semitones, $V_{nj} = 3.62$; (b) Element spacing according to a dominant-seventh chord, $V_{nj} = 8.6$; +: 2-chord, ×: 4-chord, ◇: 8-chord sequence; (c) Element spacing in quarter tones (+, $V_{nj} = 7.13$) and according to dominant-seventh chord (×, $V_{nj} = 61.9$), with 5-chord sequences.

3.5 Models

3.5.1 Voice Tracing Model

When one listens to an RCS a sequence of simultaneous tones is heard. When a subject is asked to determine the direction of global pitch motion heard in such a sequence, he may, on the one hand, make the decision on the basis of apparent movement of some "average" pitch. An "average" pitch cannot be a simple frequency average for each frame, since such an average contains no relevant information about the imposed frequency motion. The listener will have to form some list of element pairs that define a correspondence pattern between successive chords and average only over those pairs. Solving the correspondence problem is not a simple matter and will be returned to later. On the other hand, the subject might use the much simpler strategy of tracing a particular set of elements in successive frames. Musical practice indicates, for instance, that polyphonic dictation, where students have to write down the notes of several simultaneously sounding voices, is always easiest for the two extreme voices: soprano and bass. It therefore seems possible that subjects in the present experiments based their responses entirely on the perceived frequency movement of either upper or lower elements of successive chords.

Inspection of the data for two-chord RCSs of Experiment A from this point of view, and only for those RCSs that contained sinusoidal or sawtooth wave tones, revealed a definite correspondence between perceived pitch motion and physical frequency motion of the highest "on" elements of the two chords. Eighty percent of all trials showed this correspondence. For those trials where the highest tone frequencies were identical, the lowest

"on" elements (bass voice) appeared to have some influence on perceived pitch motion, but the correlation was not very large. For RCSs of four and eight chords the criterion for upward vs downward frequency motion of the top (soprano) or bottom (bass) voice is no longer clear because of the randomness in the order of successive notes. Although some criterion could possibly be defined it was not attempted because of inherent arbitrariness.

The results obtained with the Shepard tones provide perhaps the strongest evidence that simple voice-tracing models cannot account for observed behavior. In a chord of (up to) six Shepard tones one cannot really tell which tone is the highest or the lowest because of the circular pitch property exhibited by such tones. A "soprano" or "bass" voice is, even by purely physical criteria, impossible to determine. The data of Figure 3.4.1(c) however, clearly show that correlation between the directions of frequency motion (N) and perceived global pitch motion is greatest precisely for this type of chord elements. This suggests that global pitch movement is not a perceptual feature of a few particular elements in the chord sequence, but is probably a global feature to which all elements can in principle contribute.

3.5.2 Dipole Contribution Model

In order to determine which elements in an RCS contribute to the percept of pitch motion, the RCS is separated into its smallest elementary events for which pitch motion can be perceived. Every two elements (tones) in two different chords (frames) form such an elementary event. These events are called *dipoles*. A dipole D is characterized by (a) its displacement vector $\mathbf{d}=(f,t)$, where f is the number of tone (frequency) steps within the given

chord structure and t is the temporal distance expressed in frames, and (b) its form h denoting the states of the two dipole elements. Since there are only two possible states for any element, there are four possible dipole forms: on-on, on-off, off-on and off-off, denoted as h_{11}, h_{10}, h_{01} and h_{00} respectively. An example is shown in Figure 3.5.1.

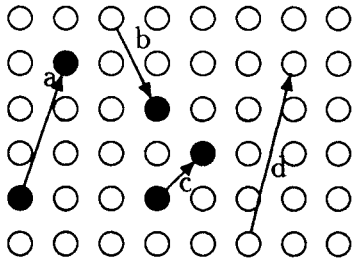


Figure 3.5.1: Examples of various types of dipoles. Dipole a: $\mathbf{d}$*=(3,1);*h_{11}. *Dipole b:* $\mathbf{d}$*=(-2,1);*h_{01}. *Dipole c:* $\mathbf{d}$*=(1,1);*h_{11}. *Dipole d:* $\mathbf{d}$*=(4,1);*h_{00}.

The main idea of the Dipole Contribution Model (DCM), proposed by Allik and Dzhafarov (1984), is that each dipole contributes in principle to perceived global pitch motion. Some dipoles, such as the ones with displacement vectors $\mathbf{d}=(0,0)$, $(0,1)$ or $(1,0)$, are excluded because they cannot possibly convey such a percept. The dipole contribution, $\mathrm{c}(\mathrm{D})=\mathrm{c}(f,t,h)$, is a unidimensional random variable representing the perceptual contribution of this elementary event to apparent pitch motion. The sign of this contribution is taken positive if the displacement vector $\mathbf{d}$ points upward (f is positive) and negative if it points downward (f is negative). The statistics of c(D) depend on the dipole's displacement vector $\mathbf{d}$ and form h. Dipole contributions c(D) are assumed to have the following general properties:

(a) Homogeneity. All dipoles of the same displacement vector $\mathbf{d}$ and form h contribute equally, regardless of their position of occurrence within a RCS.

(b) Symmetry. Contributions of any two symmetrical dipoles, i.e., dipoles of the same form h but with displacement vectors $\mathbf{d}=(f,t)$ and $\mathbf{d}=(-f,t)$, respectively, have symmetrical probability density functions so that the average of their net contribution is zero.

(c) Independence. Contributions of any two different dipoles are statistically independent, provided their types and forms are known. (When the dipole form is a random variable, as it is treated in most of the remainder of this paper, contributions of two dipoles are not always independent because of the interdependence of forms for some dipole pairs in the configuration).

The contributions of all the various dipoles are arithmetically added to form a Total Sum of Contributions (TSC) which, of course, is also a random variable. If for a given chord sequence the TSC is positive, the perceived direction of global pitch motion will be upward, and if negative, it will be downward. For any value of SRP and N the relative frequency of occurrence of the various dipoles can be determined statistically. It is assumed that, since the TSC is the sum of a large number of relatively small contributions, most of which are mutually independent, the TSC can be considered as approximately Gaussian and specified by its two characteristic parameters E[TSC] and Var[TSC]. The probability Pc of a "correct" identification of global pitch motion, i.e., a perceived motion that agrees with the actual direction of frequency transformation N, is:

$$Pc = \frac{1}{\sqrt{2\pi \mathrm{Var[TSC]}}} \int_0^\infty \exp[-\frac{(\mathrm{x} - \mathrm{E[TSC]})^2}{2\mathrm{Var[TSC]}}]\mathrm{dx}. \qquad (2)$$

This is illustrated in Figure 3.5.2, in which "correct" identification of either frequency motion direction has been indicated by the two differently-shaded areas. The integral of Eq. (2) can easily be reduced to the standard Gaus-

sian integral:

$$Pc = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{E[TSC]}{\sigma[TSC]}} \exp[-\frac{x^2}{2}]dx = \Phi(\frac{E[TSC]}{\sqrt{Var[TSC]}}), \quad (3)$$

where $\sigma[TSC] = \sqrt{Var[TSC]}$.

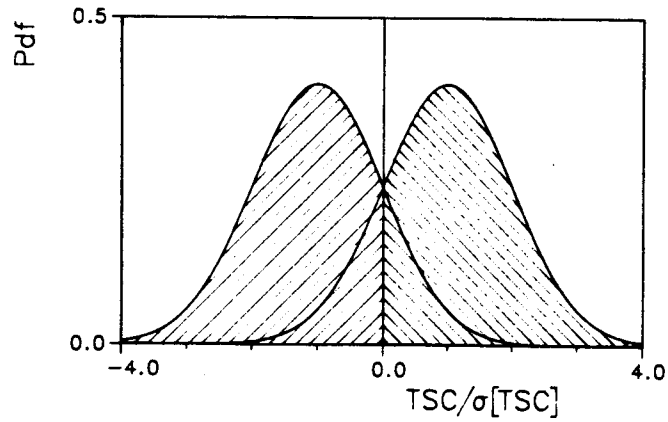


Figure 3.5.2: Gaussian detection model relating observed identification performance to statistics of the DCM. The decision criterion is assumed to be balanced, i.e., placed to maximize 'correct' identifications.

Although Eq. (3) may suggest that only two free model parameters are involved, there are, in principle, very many because the magnitudes of c(D), i.e., the details of density functions, are freely chosen for each dipole of a particular displacement vector and form. A model with so many free parameters is not testable and therefore not too interesting. To restrict the model further, three additional simplifications were made:

(1) The DCM was restricted to a *local* or *short-range* model in which only those dipoles contribute to the TSC that have a displacement vector **d** which is either (-1,1) or (1,1). This implies that global pitch motion is affected only by single-tone (frequency) jumps within successive time frames.

Dipoles that span several frequency steps and/or extend over nonsuccessive chords are ignored. This simplification was found to work well for the visual experiments on motion perception in circular random cinematograms by Allik and Dzhafarov (1984);

(2) Dipoles of the form h_{00}, i.e., off-off, do not contribute to perception of global pitch motion and are therefore ignored. This assumption is specific for the present auditory experiments because perception of pitch motion for two successive tones that are both off seems highly unlikely. Dipoles of the form h_{11}, i.e., on-on, are the only ones that convey a percept of pitch motion. They are referred to as *jumps*;

(3) Contributions by dipoles of the forms h_{01} and h_{10}, referred to as *non-jumps*, have identical distributions, with expected values equal to zero.

Because these assumptions leave only two kinds of contributions, namely those of *jump* and *non-jump* dipoles, the number of parameters of the DCM has been reduced to only two. Allik and Dzhafarov (1984) further found that their visual data could be well accounted for with a model version in which the jump contributions were a fixed, deterministic number (random variable of zero variance), leaving the variance of non-jump contributions to be estimated as the only free parameter of the model. This simplification will be considered as well.

Although the computation of model predictions for perceived pitch direction with the use of Eq. (3) is in principle straightforward, the actual computation of the necessary statistics E[TSC] and Var[TSC] for a given RCS can be very cumbersome and tedious. Because such computations have been shown in detail by Allik and Dzhafarov (1984), we will present

here only the highlights for that particular form of the DCM that applies to the present sound experiments.

The expected value of the TSC is given by the expression:

$$\mathrm{E[TSC]} = ME_{11}^{+}(\frac{P}{2} - \frac{1}{4}), \tag{4}$$

where M represents the total number of dipoles with a displacement vector $\mathbf{d}=(1,1)$ in the RCS, E_{11}^{+} is the mean perceptual contribution of dipoles with that displacement vector and form h_{11} (on–on), and P is a simple abbreviation of the state repetition probability SRP. The derivation of this equation is shown in Appendix Ia.

The variance of TSC comprises two sets of terms:

$$\mathrm{Var[TSC]} = \sum_{i} \mathrm{Var}[c(d_i)] + 2\sum_{i}\sum_{j\neq i} \mathrm{Cov}[c(d_i), c(d_j)]. \tag{5}$$

The first term, as is shown in Appendix Ib, is equal to:

$$\sum_{i} \mathrm{Var}[c(d_i)] = \frac{1}{4}M(E_{11}^{+})^2\{(2P+1)V_j + (2Q+1)V_{nj} + P + PQ + \frac{3}{4}\}, \tag{6}$$

where P is the SRP, Q equals $1 - P$, and V_j and V_{nj} are the normalized variances of contributions by dipoles in jump form and the non-jump form, respectively:

$$V_j = \frac{\mathrm{Var}[c(d_{11}^{+})] + \mathrm{Var}[c(d_{00}^{+})]}{(E_{11}^{+})^2},$$

$$V_{nj} = \frac{\mathrm{Var}[c(d_{01}^{+})] + \mathrm{Var}[c(d_{10}^{+})]}{(E_{11}^{+})^2},$$

The second term of Eq. (5) is a sum of covariances which can be written as:

$$2\sum_{1}\sum_{j\neq i} \mathrm{Cov}[c(d_i), c(d_j)] = \frac{1}{2}(E_{11}^{+})^2 \sum_{i,j\in S_1,S_2} \{P(d_i, d_j) - \frac{1}{4}\}, \tag{7}$$

where $P(d_i, d_j)$ is the joint probability that the ith and jth dipoles are both of the form h_{11} (on–on) and the summation is taken over covariances of those single-step dipole pairs (d=$(\pm 1, 1)$) that lie along a diagonal in the direction of frequency motion in the RCS (set S_1) or that connect two of those successive diagonals (set S_2). This is because the effect of the SPR propagates along those diagonals and causes a correlation between the forms of dipoles contained in or touching them. All dipoles pairs not belonging to the sets S_1 or S_2 have a joint probability $P(d_i, d_j)$ of being in the h_{11}-form that equals 0.25, which makes their covariance zero. Unfortunately there is no simple general expression for $P(d_i, d_j)$ for the dipoles contained in the sets S_1 and S_2 since such an expression depends on the length of the chord sequence. For any chord sequence, however, $P(d_i, d_j)$ ultimately depends only on the state repetition probability SRP, as is shown in Appendix Ic.

If we now return to Eq. (3), we see with the aid of Eqs. (4), (6) and (7) that its independent variable $\mathrm{E[TSC]}/\sqrt{\mathrm{Var}[TSC]}$ depends only on M, P(=SRP), V_j and V_{nj}, where the latter two are free parameters. In the following data analysis it will be assumed that the parameter V_j is equal to zero, which amounts to assuming that each observed on–on jump on a RCS causes a fixed, noiseless contribution to the overall perceived pitch motion. The only remaining model parameter V_{nj}, representing perceptual noise contributed by observed dipoles that have a non-jump form, can then be estimated by fitting Pc(SRP), plots of "correct" direction identification as a function of SRP generated by Eq. (3), to the empirically obtained data.

For each of the three subject-averaged data sets shown in Figure 3.4.1 respectively obtained with (a) sine-wave chords, (b) sawtooth-wave chords, and (c) Shepard-tone chords, three functions were computed according to Eq. (3) with the parameter V_{nj} chosen in such a way that the summed square error between function and data values was minimized. In other words, one best-fitting V_{nj}-value was computed for data from 2-chord, 4-chord and 8-chord sequences of particular sounds. Only half of the empirical data were used to estimate V_{nj} and the corresponding function. These functions are shown as solid curves in Figure 3.4.1. The same procedure was followed for the data of Experiment B, using all the data points. The results are shown in Figure 3.4.2.

A chi-square test was performed on the fitted results of Experiment A, using the second half of the empirical data. It was found that model functions and empirical data (i.e. those data that had not been used to obtain the functions !) did not differ significantly at the $p < .05$ level.

3.6 Discussion

The data of both experiments A and B are well accounted by a simple version of the Dipole Contribution Model (DCM). A simple version means that (a) it is a short-range local model in which correspondence between elements is limited to adjoining elements in successive time frames, (b) all *off-off* jumps do, on the average, not contribute (*no-blank* version), and (c) the noise involved in the perception of *on-on* jumps is assumed to be zero ($V_j = 0$). The model accounts for the reversed direction of perceived pitch motion, apparent from the data, for state repetition probabilities smaller than 0.5 (the so-called "reversed phi-phenomenon"), as well as the slight asymmetry of empirically obtained psychometric functions about the point $Pc = 0.5$, SRP=0.5. The chi-square test performed with one half of the data of Experiment A on the model-generated curves that were optimally fitted to the other half of the data yielded values that were consistent with the hypothesis that both halves of the data are based on the same underlying mechanism. The chi-square criterion did not, however, allow us to discriminate unequivocally between possible variants of the short-range local DCM, for instance assumptions that $V_j \neq V_{nj} \neq 0$ (two free parameters) or that $V_j = V_{nj} \neq 0$. The former assumption, when put into the model, yields somewhat smaller chi-square values, as is expected with two free parameters, but also typically yields optimum estimates for V_j about ten times smaller than for V_{nj}.

The implication of parameter values found for V_j and V_{nj} is that *on-on* relationships between corresponding tone elements are perceived with much less noise interference than *on-off* or *off-on* relationships. Normally

one might expect the amounts of noise to be about the same. Although a chi-square test did not provide compelling evidence to reject this hypothesis, the data appear generally more consistent with the idea of a (relatively) much smaller V_j. For reasons of simplicity we chose to represent the global pitch perception process by the rather extreme version of the DCM in which $V_j = 0$, leaving only the free parameter V_{nj} to be estimated, without the intention of claiming that the percepts of short-range *on-on* jumps are totally noiseless.

Another potentially significant finding is that perception of global pitch motion within the general context of these experiments can be well accounted for with a short-range, local form of the DCM. There remains the question, however, to what extent this finding can be generalized to all music perception. Experiments with tone streaming have convincingly shown that in sufficiently rapid tone sequences our auditory system may group tone elements together that are not in successive time frames. It is possible, for instance, to construct a single sequence of tones which leads to a percept of two parallel melodies formed by odd and even numbered notes, respectively (Bregman & Campbell, 1971; van Noorden, 1975). There are also visual analogs to those streaming experiments (Julesz & Bosch, 1966). It is quite possible that the short-range local solution to the element correspondence problem, which the auditory system seems to employ in the present experiments, has something to do with the large amount of randomness in the chord sequences. Harmonic or melodic structure, put into transitions from one chord to the next, could cause the auditory system to use solutions to the correspondence problem other than a short-range

local one. This implies that 20th century serial music, in which traditional harmonic and melodic structures play a negligible role (Austin, 1966), is perceived in a short-range and local manner by most listeners. If this conclusion is correct, a composer may have a possible tool to manipulate a listener's perceived correspondence between notes by varying the amounts of randomness and structure in the music.

The local form of DCM can also account for some auditory pitch paradoxes reported in the literature. One of them, actually used in this study, happens when a Shepard tone is followed by another Shepard tone in which all frequencies have been multiplied by 1.89. Instead of hearing the pitch go upwards a major seventh, subjects typically report hearing a pitch descent of a semitone. This is because there actually are many semitone frequency jumps in a downward direction in this case which apparently dominate the global percept. Another paradox, recently reported by Schroeder (1986) and Risset (1986), has a slightly stretched Shepard tone (e.g. 49.6, 102.4, 211.2..... Hz) followed by its exact octave transposition (i.e. 99.2, 204.8, 422.4..... Hz). Despite the component frequency jumps of exactly one octave upwards, subjects typically report hearing a slightly descending pitch jump. Apparently also in this case, *local* dipoles formed by frequencies that are close dominate the global pitch percept and are stronger than the *central* or *virtual* pitch which should have resulted in upward octave percepts.

3.7 Conclusions

Considering the quantitative behavior of the parameter V_{nj}, it is found that:

1. It appears to be rather subject-independent. For that reason data of four subjects in Experiment A and two subjects in Experiment B were pooled.

2. It seems to decrease with the harmonic richness and complexity of signals employed to represent the tone elements of the chords. The largest decrease of V_{nj}, however, occurred between the data shown in Figures 3.4.1(b) and (c), representing sawtooth and Shepard tones, respectively, whereas the change between the cases of sine and sawtooth representation (Figures 3.4.1(a) and (b) was rather small. Unfortunately, the change from sawtooth to Shepard tones involved not only a change in tonal spectra but also a change of tone spacing within chords (semitone to whole-tone spacing) and, more importantly, a change to a situation of true physical circularity of the chord sequence. This covariance of parameters obscures a clear-cut conclusion on what exactly causes the decrease in the noise of perceived non-jumps.

3. Regularity of inter-tone spacing generally leads to smaller values of V_{nj}. This can be seen by comparing results of Figures 3.4.2(a) and (b), and also the two sets of data shown in Figure 3.4.2(c). Despite the larger inter-element tonal distances of the dominant-seventh chord,

compared with semitones or quarter tones, the resulting estimates of V_{nj} are consistently and significantly larger. This is, in a sense, logical because with a regular inter-element space the processor has only to deal with a single dipole size, whereas in the dominant-seventh chord arrangement the physical frequency jumps of nominally identical dipoles are often different in size. This can only be a confounding factor for any processor, and is most likely to lead to degraded performance.

4. The auditory results presented in this study appear very similar to the visual data obtained with random circular cinematograms by Allik and Dzhafarov (1984). Both sets of data are adequately accounted for by very similar models. This suggests that in both cases we are apparently dealing with a general cognitive brain process which extracts information from visual and auditory sensory inputs in a similar manner.

4 Deterministic Signals in Discrete Noise

4.1 Introduction

This chapter deals with the perception of a deterministic signal of a discrete nature against background of randomly changing tones. In experiments with random-chord sequences (Chapter 3) a statistical arrangement of tones causes a global pitch motion percept. The stimuli in the experiments in this chapter resemble the stimuli from chapter 3 very much (in some cases they can be identical), but the subject's task is fundamentally different. In experiments with RCSs the subject is asked to determine the global pitch motion, whereas in this experiment the subject is asked to identify or separate the signal from the background noise tones. The signal is an either ascending or descending musical tone scale. To our knowledge, no experiments are reported where a melody is presented against a background of other potentially interfering tones. One of the few situations that slightly resemble our experiments is the dichotic conflict situation in which two different melodies are simultaneously presented to a subject, one to each ear (Kimura, 1964; Deutsch, 1975; Butler, 1979). The limits of our auditory system to perceive a melody against a background of other potentially interfering tones without dichotic separation remains largely unknown. This is somewhat unfortunate because it is just this situation that is encountered most often when we listen to music.

In research on vision, the problem of motion perception for discretely changing elements against a background of other randomly changing el-

ements seems to have been given more attention that the above mentioned analog in hearing. Experiments with moving random-dot patterns or *cinematograms* (Julesz, 1971; van Doorn and Koenderink, 1982, 1984; Nakayama and Silverman, 1984) have convincingly shown that, when a directional moving dot signal is embedded in a background of randomly moving dots, the motion can often be detected. This implies that the brain must have a rather sophisticated way of solving the correspondence problem.

The experiments presented in this chapter are designed to get some insight in the detection process and the strategy used by subjects by listening to complex signals. Also will be studied to what degree noise influences the detectability of the signal.

4.2 Experiments

In this section two experiments are performed. The first one deals with signals in overall noise (Experiment A), the second one (Experiment B) with noise, occuring only in elements of the alternative signal.

4.2.1 Stimuli

The structure of the stimuli used in this experiment can be represented by a 6 × 6-matrix, where each column represents a time frame and each row represents a tone (see Figure 4.2.1). Throughout the experiment the 6 tones were sinusoids with frequency of 262, 294, 330, 369, 415 and 466 Hz, representing the six tones of a whole-tone scale from C to B♭ in equal temperament.

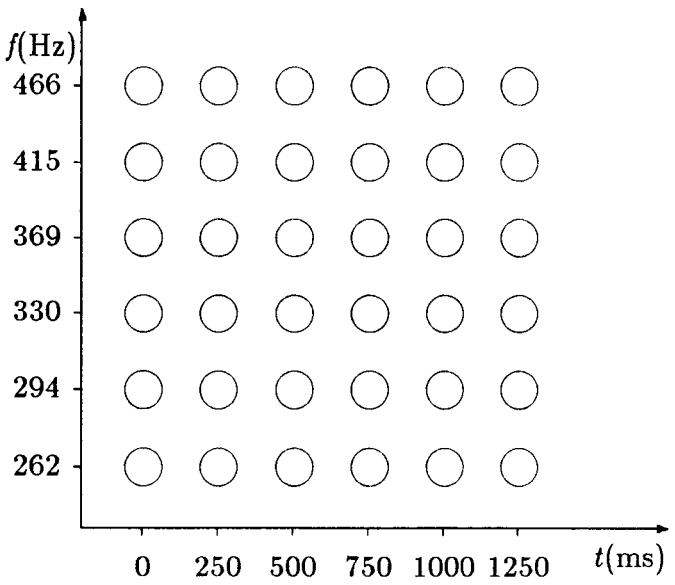


Figure 4.2.1 : Representation of a stimulus. The abscissa is time; the ordinate is frequency.

The duration of one time frame was 250ms, which included a 20-ms on and

20-ms off ramp. There were no silent periods between tones. One tone from one time frame is called an *element*. One stimulus contains 36 elements. Each element can be put to state *on* (audible) or *off* (silent) independently. All *on*-elements have equal amplitude. Sometimes an *on*-element is called a 1-element, an *off*-element a 0-element. Two elements lying in successive time frames and have adjacent frequencies form a *dipole*. If e.g. both elements are 1-elements the dipole is called a 11-dipole.

Each stimulus contained one of two possible signals (i.e. a sequence of 1-elements, which had to be identified) plus noise-elements. The first signal was an ascending whole tone scale (see Figure 4.2.2(a)). This signal was labeled *up*. The second signal was a descending whole tone scale (see Figure 4.2.2(b)). This signal was labeled *down*.

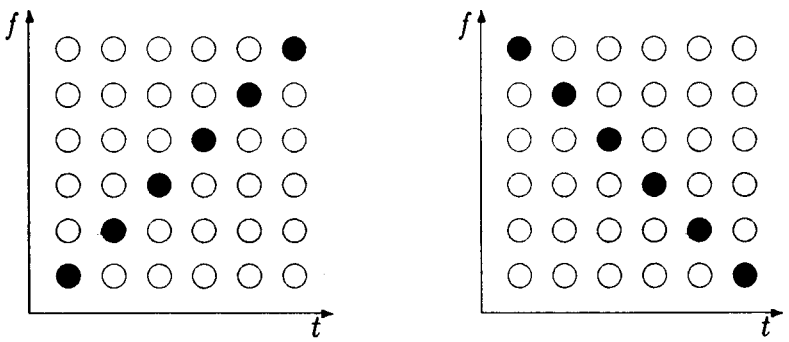


Figure 4.2.2 : (a) Representation of an up-stimulus. (b) Representation of a down-stimulus. An open circle denotes a 0-element; a closed circle a 1-element.

Notice that both signals lie at the diagonals of the matrix. We therefore sometimes speak about *the* diagonal (that diagonal where the signal is placed) and the *alternative* diagonal (the other diagonal). Elements, not

lying at one of the diagonals are referred to as *non-diagonal* elements.

In Experiment A, in order to mask the signal, elements that were not part of the signal were put *on* with a certain probability P_n. Before a stimulus was generated, a choice was made between *up* and *down* and P_n. Thus if $P_n = 0.0$, only stimuli like the ones pictured in Figure 4.2.2 were obtained. If e.g. $P_n = 0.3$ stimuli like the one in Figure 4.2.3 are obtained.

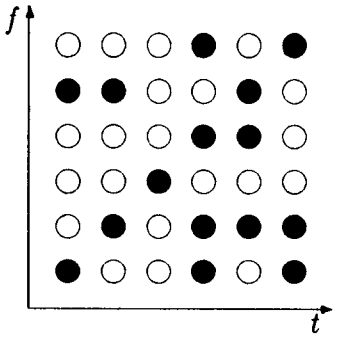


Figure 4.2.3 : Representation of a stimulus with $P_n = 0.3$.

If $P_n = 1.0$ all elements are *on* and, since both signals are physical present or, in other words, both diagonals contain only 1-elements, an *up*-stimulus is indistinguishable from a *down*-stimulus. In Experiment A P_n varied from 0.0 to 1.0 in steps of 0.1.

Of particular interest are the elements lying at the alternative diagonal. When an increasing number of 1-elements is present at the alternative diagonal, identification of the signal becomes more difficult, regardless the amount of non-diagonal elements in state *on*. To examine the importance of the elements at the alternative diagonal Experiment B was designed.

The stimuli of Experiment B resembled the ones from Experiment A. Again the stimulus contained either an *up*- or *down*-signal. However, now only elements at the alternative diagonal could be put *on*. All non-diagonal

elements remained *off* during Experiment B. Thus if e.g. the stimulus was labeled *up*, apart from the signal-elements, only 1-elements could appear at the alternative diagonal, i.e. the diagonal from top-left to bottom-right (see Figure 4.2.4). Since now there are only 64 (= 2^6) ways of arranging these elements at the alternative diagonal, we don't need P_n any more as an indicator for the noise level.

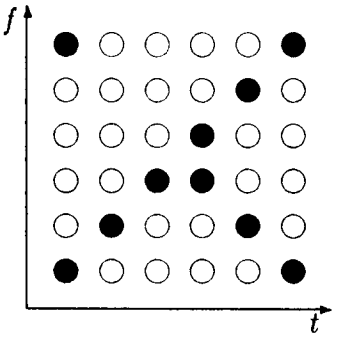


Figure 4.2.4 : Representation of an up-stimulus of Experiment B.

4.2.2 Method

All sessions started with the determination of the subject's hearing threshold for the stimulus to be used. The subject, who was seated in a double-walled (IAC) sound insulated chamber, received the stimuli binaurally through Sennheiser HD 424 headphones and adjusted the intensity of a gated 440-Hz sinusoidal tone to detection threshold. During the experiment all stimuli were presented 20 dB above this empirically established threshold. The rather low level was chosen to avoid as much possible confounding effects of aural combination tones (Zwicker, 1955; Goldstein, 1967).

After the presentation of each stimulus the subject's task was to indicate whether the signal was *up* or *down*, corresponding with the ascending and descending scale respectively. He did this by pressing one of two buttons

on a response-box. There was no response time limit, and feedback was provided immediately after each response.

For Experiment A, about 4000 trials were collected from each subject, most of them between $0.4 \leq P_n \leq 0.8$. One session contianed up to six runs of 110 trial, and lasted about one hour.

For Experiment B, about 1000 trials were collected from each subject, all of them having at least three 1-elements at the alternative diagonal.

4.2.3 Subjects

All trials in this experiment were judged by four subjects: AH, JG, RL and NV. The degree of musical experience as well as the experience in psychoacoustical tasks, however differed greatly from subject to subject. All subjects had normal hearing.

4.2.4 Apparatus

The stimuli were calculated on a P857 mini-computer system at a sampling rate of 10 kHz. With a 12-bit D/A-converter these data were transformed into acoustical signals. The output of the D/A-converter was low-pass filtered at 4.3 kHz with two Krohn-Hite 343 filters (96dB/octave). All signals previously have been checked with a Data Precision Analyser (DATA-6000). The frequency-response curve of the headphones appeared to be flat within 3.5 dB in a region from 60 to 5000 Hz.

4.3 Results

This section deals with the results of Experiment A and B. In the figures presented in this section, the subject's score is presented in units of d' (as described in section 2.1). d' is related to the %-correct score following

$$P_c = \Phi(d'/2),$$

where $\Phi(z)$ is the Normal Probability function. Table 4.3.1 shows some values for d' with its associated %-correct score.

d'	P_c(%)	d'	P_c(%)
0.0	50.0	3.0	93.3
1.0	69.1	4.0	97.7
2.0	84.1	5.0	99.4

Table 4.3.1 : Relation between d' and P_c.

The data in the figures in this section are pooled over all subjects, since inter-subject consistency prooved to be good. The (numerical) results for the individual subjects can be found in Appendix II.

4.3.1 Experiment A

Figure 4.3.1 shows the results of Experiment A where d' is plotted versus the noise level P_n. $P_n = 0.0$ means that no elements are *on* except the ones which belong to the signal. $P_n = 1.0$ means that all elements are *on*. Detection then is at chance level. $P_n = 0.3$, e.g., means that, on the average, 3 out of 10 elements are in state *on*.

When $0.5 \leq P_n \leq 1.0$, d' seems to be a linear function of P_n. If the line

$l : d' = a + b(1 - P_n)$ is fitted to this part of the data of Figure 4.3.1, a value for the offset a is obtained, lying close to zero. Since (1) for $P_n = 1.0$ d' has to be zero (all elements are *on*); (2) the standard error of a allows us to put a to zero; (3) there is no reason to assume that, apart from the noise caused by P_n, there exists other noise or signal uncertainty which might cause an offset a, we take $a = 0$, and fit the line $l : d' = b(1 - P_n)$ to the data. This gives $b = 3.7$.

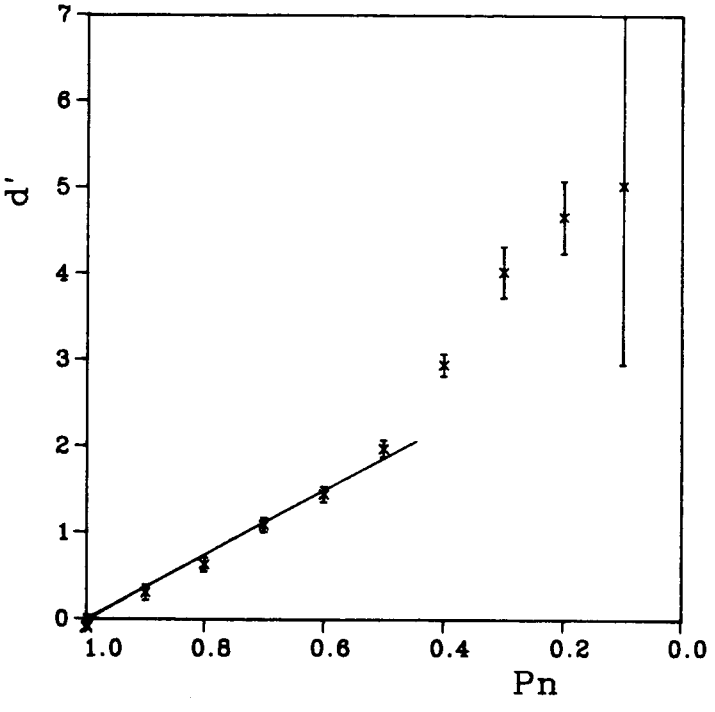


Figure 4.3.1: Results of Experiment A. d' is plotted versus the noise-probability P_n.

If $0.0 \leq P_n \leq 0.4$ the theory as described in section 2.1 cannot be properly applied since (1) the number of trials per data point for $0.0 \leq P_n \leq 0.4$ is small and (2) d' is rather large, so the identification task is so easy that an almost 100% score is obtained. Therefore the data points for $0.0 \leq P_n \leq 0.4$

are not taken into account in the fit. For the sake of completeness these points are presented in Figure 4.3.1. When $P_n = 0.0$ all subjects scored without one mistake, therefore the associated d' is infinetely large.

Figure 4.3.2 shows d' versus the number of 1-elements N at the alternative diagonal. If N=0, all elements at the alternative diagonal were 0-elements; if N=6, all elements at the alternative diagonal were 1-elements, and, since both signals then are present, performance was at chance level.

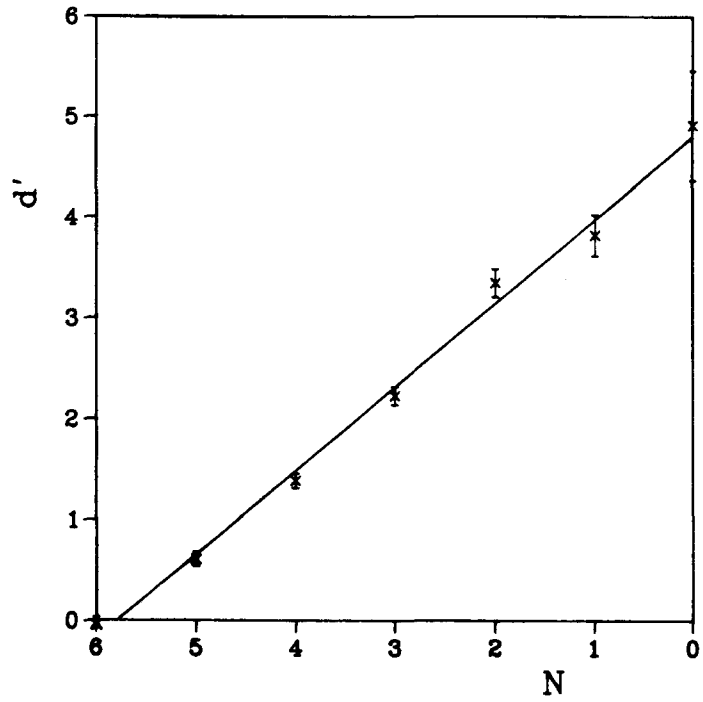


Figure 4.3.2: Results of Experiment A. d' is plotted versus N, the number of 1-elements at the alternative diagonal.

If the line $l : d' = a + b(6 - \mathrm{N})$ is fitted to the data, it appears that a is significantly non-zero and l intersects with the N-axis for N=5.80. This offset probably is due to the variation in noise at non-diagonal elements. Furthermore $b = 0.83$.

One may wonder, since N is an integer, if it is allowed to speak about N=5.80 and to fit a line. When one keeps in mind that we deal with experiments with a 6 × 6-matrix, we can take N only equal to 1, 2, 3, 4, 5 or 6. If, on the other hand, we dealt with a 5 × 5-, 7 × 7- or $n \times n$-matrix an analogue experiment could be performed. Then probably again a linear relationship between N and d' would be obtained. If we plotted d' versus N/5 (in case of the 5 × 5-matrix), versus N/6 (in case of the 6 × 6-matrix), versus N/7 (in case of the 7 × 7-matrix), or versus N/n (in case of the $n \times n$-matrix), one might expect that in each case about the same curve would be obtained. Thus actually we should speak about the fraction N/6 of elements being on at the alternative diagonal, but for convenience we choose to work with N. This reasoning is not entirely correct, since we don't know the subject's performance for $n = 5$ or 7. Moreover, later in this section we will see that the first and last elements at the alternative diagonal are more important for a correct identification than the middle elements.

Figure 4.3.3 shows d' versus the number of 11-dipoles D at the alternative diagonal. D=0 corresponds with no 11-dipoles (which does not mean that there are no 1-elements at the alternative diagonal); D=5 corresponds with 5 11-dipoles, thus 6 1-elements. At D=5 $d' = 0$, since both signals are present in the stimulus.

With 6 elements at the alternative diagonal D=4 can be constructed in only two ways: 111110 or 011111 (where an 1 denotes an *on*- and a 0 an *off*-element). D=4 thus represents the score where either only the first or the last element was *off*. Figure 4.3.3 shows that identification for D=4 is

good, significantly better for D=3 and equally good for D=2 and D=1.

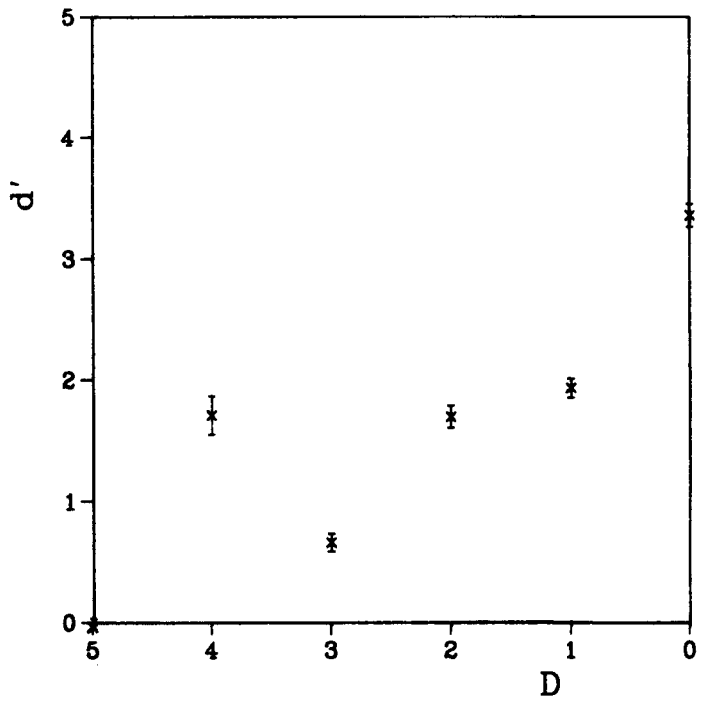


Figure 4.3.3: Results of Experiment A. d' is plotted versus D, the number of 11-dipoles at the alternative diagonal.

The score belonging to D=3 is rather bad. D=3 contains mostly 0-elements in the middle and 1-elements at the ends, e.g., 110111.

Closer examination of the data shows that 62% of all mistakes occurs with stimuli where the alternative diagonal is of type 1xxxx1 (x can be 1 or 0), the 111111-stimuli excluded; 23% of type 0xxxx1; 13% of type 1xxxx0 and only 3% of type 0xxxx0.

Figure 4.3.4 shows d' versus ΣD, the net number of dipoles in the entire stimulus.ΣD is defined as the number of 11-dipoles *in* the direction of the signal *minus* the number of dipoles in the *opposite* direction. Thus, e.g., if the stimulus is *up* and there are 7 11-dipoles (5 11-dipoles from the signal

included) in the *up*-direction and 3 in the *down*-direction, $\Sigma D = 4$.

Figure 4.3.4 shows that by decreasing ΣD d' decreases to $\Sigma D \approx 0$. If $\Sigma D = 0$ $d' > 0$ and by a further decrease of ΣD d' slightly increases.

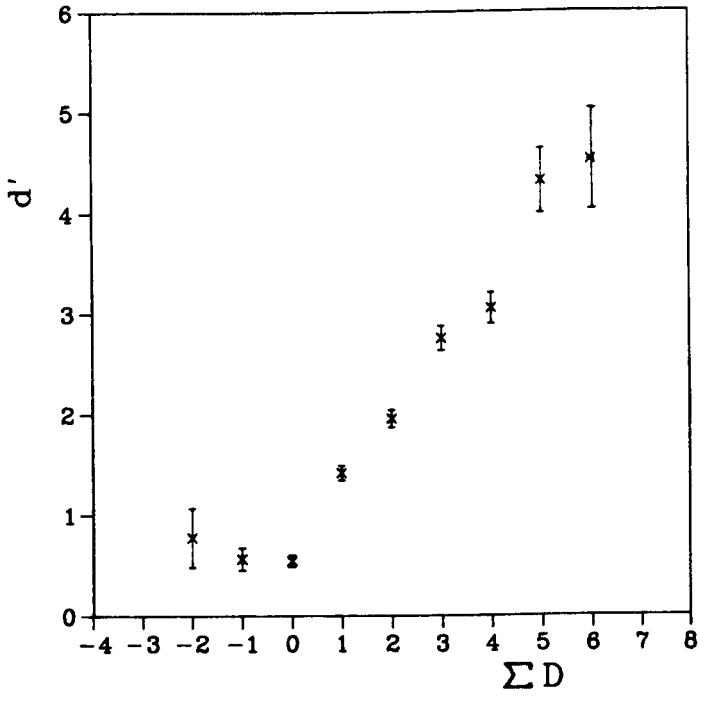


Figure 4.3.4: Results of Experiment A. d' is plotted versus ΣD*, the net number of 11-dipoles in the direction of the signal.*

4.3.2 Experiment B

Figure 4.3.5 shows the results of Experiment B where d' is plotted versus the numbers of 1-elements N at the alternative diagonal. If N=0, all elements at the alternative diagonal were 0-elements; if N=6, all elements at the alternative diagonal were 1-elements, and, since both signals then are present, performance was at chance level.

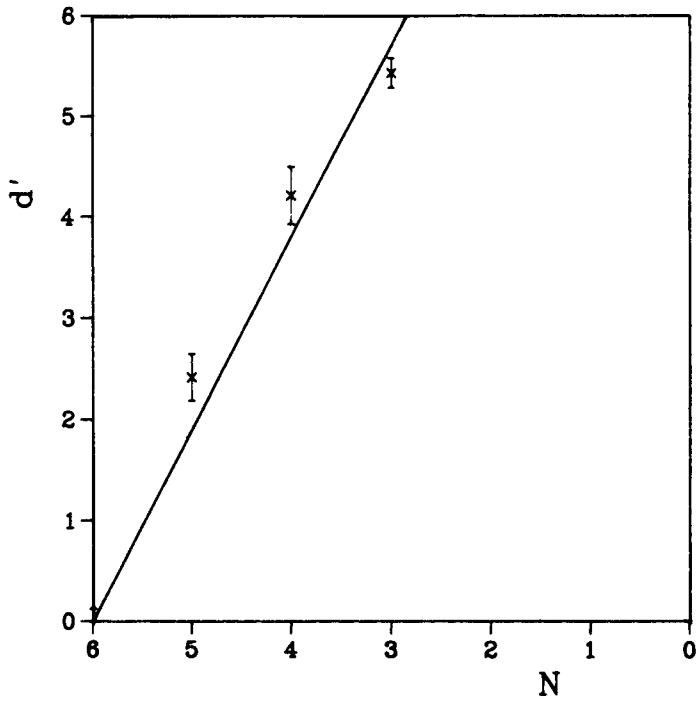


Figure 4.3.5: Results of Experiment B. d' is plotted versus N, the number of 1-elements at the alternative diagonal.

Fitting the line $l : d' = a + b(6 - \mathrm{N})$ to the data would give $a > 0$. This is not possible since, at N=6 performance should be at chance level. Therefore and because the standard error allows it, we put $a = 0$. The line $l : d' = b(6 - \mathrm{N})$ then gives us $b = 1.9$.

Figure 4.3.6 shows d' versus the number of 11-dipoles D at the alternative diagonal. D=0 corresponds with no 11-dipoles (which does not mean that there are no 1-elements at the alternative diagonal); D=5 corresponds with 5 11-dipoles, thus 6 1-elements. At D=5 $d' = 0$ since both signals are present in the stimulus.

As in Figure 4.3.3 performance for D=4 (stimuli where only either the first or the last element was a 0-element) is better than for D=3.

Closer examination of the data shows that 80% of all mistakes occurs with stimuli of type 1xxxx1 (x can be 1 or 0), the 111111-stimuli excluded; 15% of type 0xxxx1; 4% of type 1xxxx0 and only 1% of type 0xxxx0.

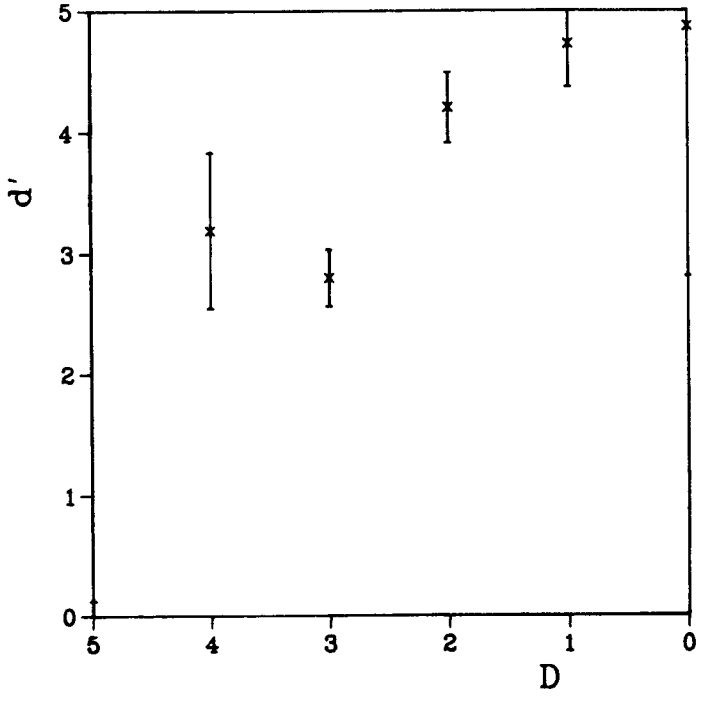


Figure 4.3.6: Results of Experiment B. d' is plotted versus D, the number of 11-dipoles at the alternative diagonal.

4.4 Discussion

The aim of the experiments was to discover which strategy subjects develop to separate the signal from the background. An ideal subject would adopt the strategy of tracing one of both possible signals, say *up*. Thus at the beginning of a stimulus, he focuses at the lowest element and find it's a 1-element. The next time frame he focuses at the one but lowest element and find it's a 1-element, et cetera, till he finds a 0-element. Then the ideal subject knows that the signal was not *up* and responds *down*. If he finds only 1-elements at this diagonal, he knows that can be *up* and responds *up*.

A human subject tries to adopt this strategy, but doesn't control it perfectly. He will be influenced to some degree by other 1-elements, either at the alternative diagonal or even by 1-elements at other places. The fact that stimuli with no interfering elements always were judged correctly, showed that subjects were very well aquainted with its shape.

Comparing the results of Experiments A and B shows that the offset in Figure 4.3.2 apparently is caused by the presence of non-diagonal 1-elements, for this offset disappears when all non-diagonal elements are put *off*. Probably the non-diagonal 1-elements hamper the subject in his effort to identify the tones from the signal. The non-diagonal elements cause apart from the offset, additional noise in the identification process, since the slope of the line in Figure 4.3.5 ($b = 1.9$) is about twice as large as the slope in Figure 4.3.2 ($b = 0.83$).

The results as plotted in Figures 4.3.3 and 4.3.6 can be explained by stating that the states of first and the last elements at the alternative diagonal are more important in the identification process than the states of

the middle elements. We can deduce this from the fact that in both figures the signal for D=4 (having the states 111110 or 011111 at the alternative diagonal) is relatively good separable from the noise, whereas this is certainly not the case for D=3 (having e.g. the states 101111 (5 1-elements) or 001111 (4 1-elements) at the alternative diagonal). In both cases the most mistakes are made with states 1xxxx1 (111111 excluded) at the alternative diagonal, then 0xxxx1, then 1xxxx0 and almost no mistakes were made when the alternative diagonal was of type 0xxxx0. From this we may conclude that the state of the last element at the alternative diagonal is even more important than the state of the first one. This can be explained by assuming that in the first time frame the subject is not able to identify all elements immediately, whereas this is easier in the last time frame, since the five time frames before this one have shown the exact place of all elements. Of course the performance in Experiment B is better than in Experiment A, due to the absence of non-diagonal elements, but the global behaviour is the same.

We assume that a subject always tries to maintain the same clue, i.e. to give the answer which corresponds with the direction of the signal. If this is actually happening and one believes that the decision process is governed by the amount and direction of 11-dipoles in a stimulus, as we did in the experiments with Random-Chord Sequences, one should expect $d' < 0$ by $\Sigma D < 0$ and $d' = 0$ by $\Sigma D = 0$. Figure 4.3.4 shows that this is obviously not the case. The decrease in performance by decreasing ΣD is not due to the decrease of ΣD, but to the increase in the number of 1-elements in the stimulus. At $\Sigma D = 0$ it still is possible that there are 0-elements

at the alternative diagonal, which causes $d' > 0$. Likewise the increase in performance by decreasing ΣD (for $\Sigma D < 0$) is not due to the decrease of ΣD, but to the increase in the number of 0-elements in the stimulus, which is necessary to make many 11-dipoles in the opposite direction.

4.5 Conclusions

The results of Experiments A and B allow us to draw the following conclusions:

1. A subject, when listening to one of two possible known signals against a background of noise, adopts the strategy of tracking one of both possible signals. If he notices that the possible signal he is tracking is not complete, i.e. that the diagonal contains an *off*-element, he knows that the other possibility contained the real signal. If on the other hand, he doesn't notice any *off*-elements he assumes that he tracked the real signal.

2. The data show that subjects are inclined to track a *down*-stimulus, probably because the first element of this signal is the most upper note of the signal and therefore more easily to detect.

3. It appears that when the first or especially the last element is *off* identification of the signal becomes rather easy. The state of the last element plays a more important role than tha first one because the subject not always is able to immediately identify all elements at the beginning of a stimulus, whereas at the end he indeed can identify them.

4. An ideal observer would, when tracking one diagonal, not pay attention at any other element. Clearly a human observer is not ideal, since the presence of elements at the alternative diagonal and even non-diagonal elements decrease the score considerably.

5. From experiments with Random-Chord Sequences one might expect that the subject's decision depends on the net number of *on-on*-dipoles present at both diagonals, or even on the net number of *on-on*-dipoles in the whole stimulus. This experiment shows that this is definitely not the case, since the score does not increase by an increase in dipoles. Only the presence of individual elements and their positions in the stimulus affects the score.

5 Discrimination of Spectral Profiles

5.1 Introduction

In this chapter we deal with the detectability of changes in amplitude in 20-component complexes. In literature many experiments have been reported on the threshold level of changes in a multitone complex. Thresholds always were determined with 2-alternative forced choice experiments where in one time frame a stimulus was presented containing a multitone complex without a signal and the other time frame containing the same multitone complex with a signal (i.e. an increment in amplitude in one or more components). Thresholds were determined as a function of the interstimulus interval, the number of components in the complex, the spacing of the components over the frequency-axis (Green *et al.*, 1983), dichotic or diotic presentation (Green & Kidd, 1983), the component number of the complex (Green & Mason, 1985, Green *et al.*, 1987) frequency and phase (Green & Mason, 1985) by increments in amplitude of only one component. In these experiments the component number remained fixed during one session. In other experiments the component number varied randomly during one session (Spiegel *et al.*, 1981; Spiegel & Green, 1982). In all experiments the authors tried to detect the listeners' ability to detect changes in spectral shape or *profile*. An important theoretical issue was whether the listeners' sensitivity to changes of one component in more complex spectra is better than for changes in amplitude in a single sinusoid. To avoid that listeners would focuss on changes in intensity instead of profile changes, the overall intensity of the two stimuli were varied randomly in all experiments.

The experiment reported in this chapter also deals with profile analysis, but the experimental set-up is designed such that listeners are not able to focus on changes in intensity between the two stimuli any more. In the experiments described in this chapter both stimuli contain a spectral component which has been slightly varied in intensity compared to the others, but the place of this component is not the same. This causes a jump percept. By measuring the listeners' ability to identify the direction of the jump (i.e. the frequency of the incremented component becomes higher or lower) we can learn something about the listeners' ability to identify changes is spectral shape. Furthermore we investigated the perception of intensity decrements in components (instead of increments, as is done in many experiments).

5.2 Experiments

5.2.1 Stimuli

In the experiments, described in this chapter, all stimuli contained a 20-component complex with either an equal linear or equal logarithmic spacing between adjacent components (see Figure 5.2.1). The latter stimuli next are

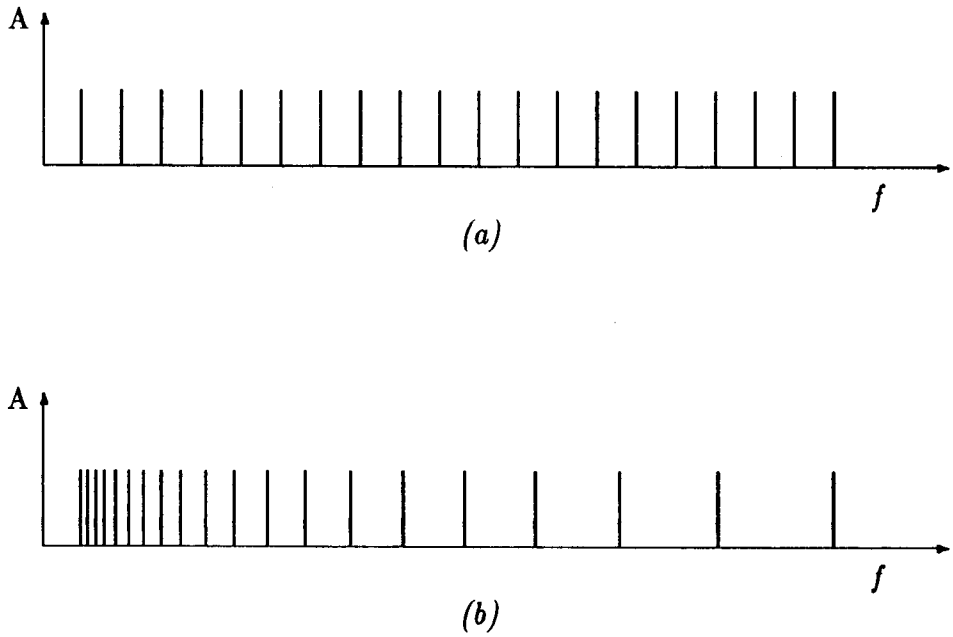


Figure 5.2.1 : Spacing of the components over the frequency-axis. (a) lin-complex; (b) log-complex.

referred to as *log*-stimuli; the other ones as *lin*-stimuli. The lowest frequency always was 200 Hz, the highest 4000 Hz, so the *lin*-complex comprised the frequencies

$$f_{lin,N} = 200 \cdot N \quad \text{Hz},$$

the *log*-complex comprised the frequencies

$$f_{log,N} = 200 \cdot 20^{\frac{N-1}{19}} \quad \text{Hz},$$

where $1 \leq N \leq 20$. Each component started in phase and all components had the same amplitude A_0. The stimulus consisted of two time frames, the duration of each time frame was 250ms which included a 20-ms on and 20-ms off ramp. There was no inter-stimulus silent period between the two time frames.

In the first time frame the N_1^{th} component was incremented from amplitude A_0 to A. In the second time frame the N_2^{th} component was incremented to the same amount, thus from A_0 to A. N_2 always was one of the neighbour components of N_1, thus $N_2 = N_1 \pm 1$ (see Figure 5.2.2).

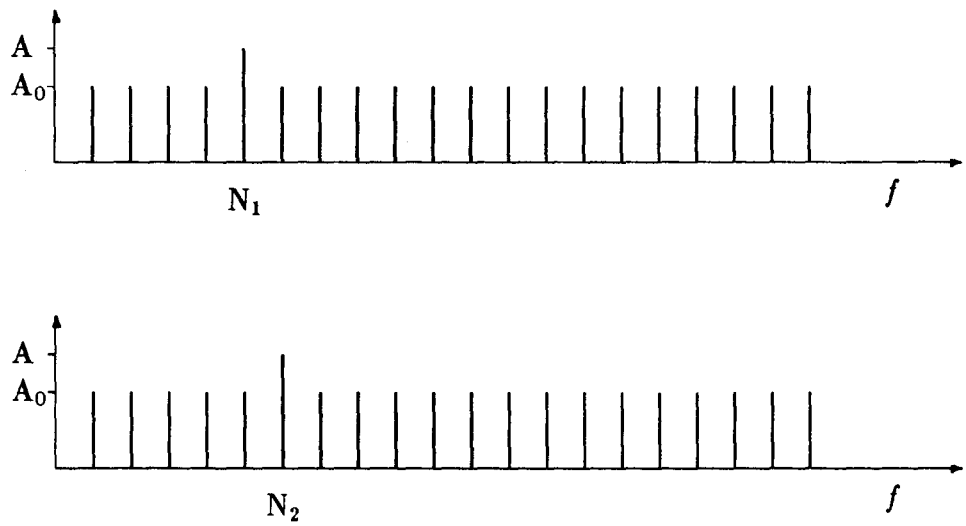


Figure 5.2.2: Representation of an up-stimulus with a lin-complex. The upper graph represents the first time frame, the lower graph the second time fame.

If $N_2 = N_1 - 1$, the stimulus was labeled *down*, due to the downward step

(or *jump*) in frequency of the incremented component from the first to the second time frame. Likewise, if $N_2 = N_1 + 1$, the stimulus was labeled *up*. A stimulus is labeled with a value N, which stands for the incremented component with the lowest frequency.

The increment in amplitude I was expressed in dB-units: $I=20\log_{10}(A/A_0)$. If I=0, there is no increment (thus no signal). If e.g. I=6.02, the increment is as large as the amplitude of the complex. Of course we also might decrease the amplitude of a component in the complex. This gives $A/A_0 < 1$, thus $I< 0$. The removal of a component results in $I=-\infty$.

5.2.2 Method

All sessions started with the determination of the subject's hearing threshold for the stimulus to be used. The subject, who was seated in a double-walled (IAC) sound insulated chamber and received the stimuli binaurally through Sennheiser HD 424 headphones, adjusted the intensity of a gated 20-component *lin-* or *log*-complex (depending at the stimuli to be used in the session) to detection threshold in about 30dB white noise. All stimuli during the experiment were presented 20dB above this empirically established threshold. The 30dB noise floor was used, in the first place to flatten the threshold as function of the frequency, so that all components of the complex would sound equally loud, in the second place, to mask as much as possible confounding effects of aural combination tones (Zwicker,1955, Goldstein, 1967). The white noise was present throughout the whole experiment, so not only when the stimulus was present.

After the presentation of each stimulus the subject's task was to indicate

whether the signal was *up* or *down*, by pressing one of two buttons on a response box. There was no response limit and feedback was provided immediately after each response.

In Experiment A about 300 to 400 trials per component per complex were collected from each subject. In this experiment N varied random from 1 to 19. The subject did not know a priori the value of N, or, in other words, where the jump would occur (*random signal-fixed mask*, see Spiegel *et al.*, 1981; 1982). One session contained up to six runs of 190 trials and lasted about one hour. In one run the increment I remained fixed (Method of Fixed Levels (section 2.2)), with values of 1.0, 1.5, 2.0, 2.5, 3.0, 4.0 and 5.0 dB.

In Experiment B about 400 trails were collected for N=4 and N=17 for the *log*-complex from subjects AH and NV. Now N remained fixed (*fixed signal-fixed mask*). During one run the increment I remained fixed at the values of 0.5, 1.0, 1.5, 2.0, 2.5 and 3.0 dB for I> 0 and -0.5, -1.0, -1.5, -2.0, -3.0, -6.0, and $-\infty$ dB for I< 0.

In Experiment C the 70.7%-correct level for I was obtained in an adaptive test for N=4 and N=17 for the *log*-complex from subjects AH and NV for I> 0 and I< 0, the procedure described in section 2.3. Trials were run in blocks of 100 and each run produced about 14 to 22 reversals, the first four reversals excluded. The increment I was decreased by 0.25 dB following two correct responses and increased by 0.25 dB following one incorrect response. The level, belonging to each configuration was determined only once.

5.2.3 Subjects

The trials in Experiment A were judged by subjects AH, JG, JS and NV in case of the *lin*-complex; and by subjects AH, RL, JS and NV in case of the *log*-complex. Trials in Experiment B and C were judged by AH and NV. The level of musical experience and experience in psychoacoustical tasks differed greatly from subject to subject. All subjects had normal hearing.

5.2.4 Apparatus

The stimuli were calculated on a P857 mini-computer system at a sampling rate of 10 kHz. With a 12-bit D/A-converter these data were transformed into acoustical signals. The output of the D/A-converter was low-pass filtered at 4.3 kHz with two Krohn-Hite 343 filters (96dB/octave). After this white noise, produced by a Wandel und Golterman Noise Generator RG-1, was added. The intensities of the stimulus and the noise could, by means of attenuators, be adjusted independently. All signals previously have been checked with a Data Precision Analyser (DATA-6000). The spectrum of the complex was found to be flat within 0.5 dB up to 2 kHz and decreased smoothly down 5 dB at 4 kHz. The frequency-response curve of the headphones appeared to be flat within 3.5 dB in a region of 60 to 5000 Hz.

5.3 Results

The data shown in the figures in this section are pooled over all subjects, since intersubject consistency prooved to be good. The (numerical) results for the individual subjects are presented in Appendix III.

5.3.1 Experiment A

Figure 5.3.1 shows the results of Experiment A, where d' is plotted versus the increment I of the random component in the complex (in dB). The diamonds represent the results of the experiments with the *log*-complex; the crosses with the *lin*-complex. The results with the *lin*-complex show a rather linear relationship between d' and I. Fitting a line to these data results in a slope b=0.873 dB^{-1} and an offset r=0.65 dB. The results obtained with the *log*-complex show for I> 2.0 a performance which increases slower than linear. Probably this is due to lack of concentration by the subjects at higher levels of I, since the task there becomes rather easy (score about 95% correct). Moreover, closer examination of the data show that some learning-effects are incorporated in these data. Therefore only the first three data points are used to fit a line. This resulted in a slope b=1.618 dB^{-1} and an offset r=0.55 dB.

The next figures show the results for each individual component N in the complex. For each value of N a least-square fit was made, which resulted in a value for the offset r, the slope b and the point $I_{70.7}$, the intensity corresponding with d'=1.0889 (the 70.7%-correct point). In case of the *log*-complex only the data points of I=1.0, 1.5 or 2.0 dB were taken into

account, for the same reason as mentioned above.

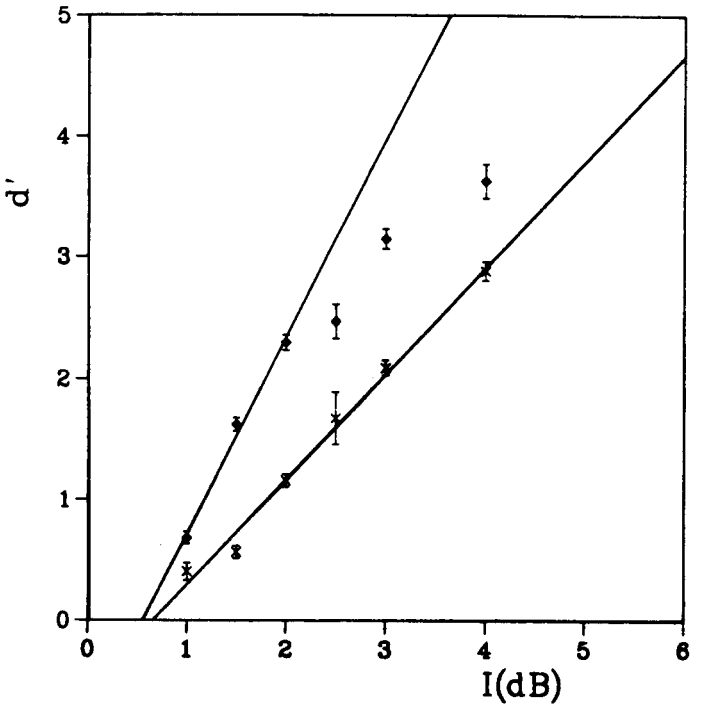


Figure 5.3.1: Results of Experiment A. d' is plotted versus the increment I(in dB) of one component in the lin(×)- and log(◇)-complex.

Figure 5.3.2(a) shows for each individual component N the slope *b* for the *log*-complex. The results show a decrease for *d'* by components near the end of the complex. Figure 5.3.2(b) shows for each individual component N the slope *b* for the *lin*-complex. The results show (as in Figure 5.3.2(a)) a decrease for *d'* by components near the end of the complex, however to a much lesser degree. For N≈ 10 a decrease in slope might be noticed. N=10 corresponds with incremented components of frequencies 2000 and 2200 Hz.

Figure 5.3.3 shows for each individual component N the offset *r* (denoted with a diamond) and the $I_{70.7}$-level *t* (for *t*hreshold). These points are

denoted with crosses.

Figure 5.3.3(a) are the results for the *log*-complex. The quantity t remains rather constant as a function of the component number N and only increases at the ends of the component range (N=1,2 and N=18,19). The offset r also remains rather constant (except for N=19). Figure 5.3.3(b) are the results for the *lin*-complex. For N=1 to N=4 both t and r decrease. For N> 4 t seems to increase slightly; r remains about constant.

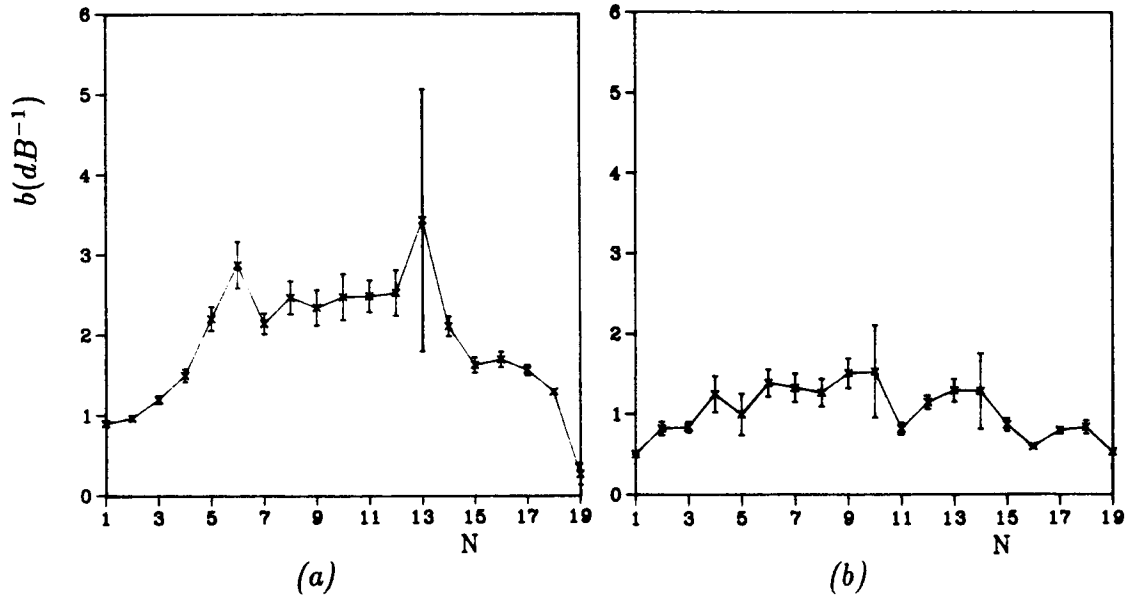


Figure 5.3.2: Results of Experiment A. The slope b (dB^{-1}) is plotted versus the component number N. (a) log-complex, (b) lin-complex.

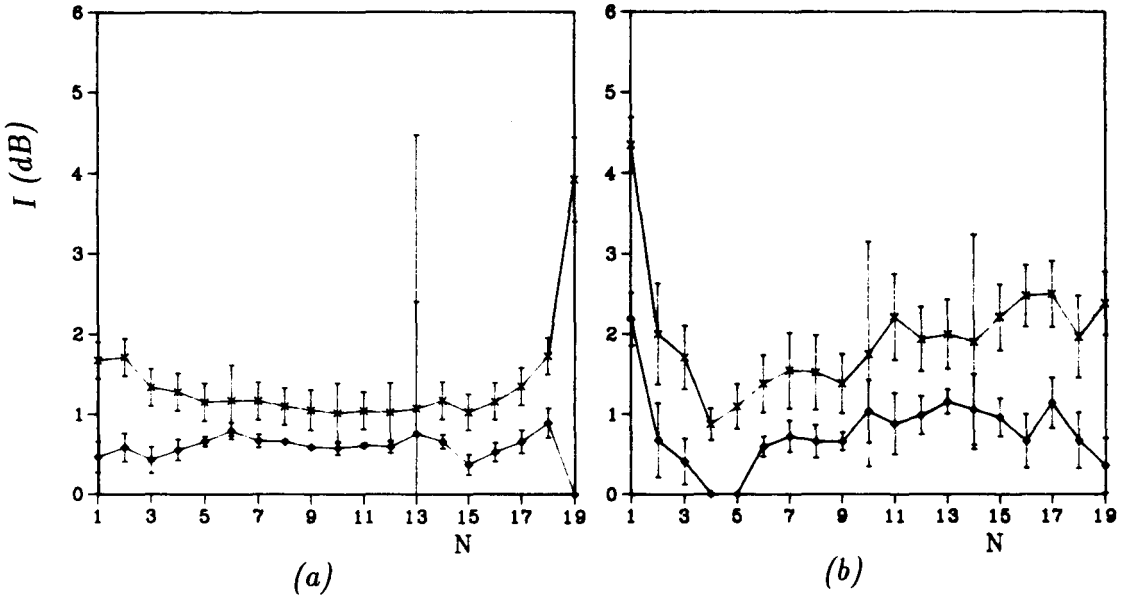


Figure 5.3.3: Results of Experiment A. the offset r (◇) and the 70.7%-correct point t (×) is plotted versus the component number N. (a) log-complex, (b) lin-complex.

5.3.2 Experiment B and C

The results of Experiment B and C are presented in Table 5.3.1, together with some relevant data from Experiment A. The experiments are performed with N=4 and N=17, for I> 0 and I< 0 (denoted with a "+" and a "–" respectively). Since in Experiment C only one point at the d'-I-plane is determined, the offset r has been put equal to zero.

Experiment A					
Subject	N	I	t	r	b
AH	4	+	1.364	0.885	2.275
AH	17	+	1.179	0.651	2.062
NV	4	+	2.742	0.000	0.397
NV	17	+	1.593	1.033	1.944

Experiment B					
Subject	N	I	t	r	b
AH	4	+	0.678	0.000	1.607
AH	17	+	1.655	1.364	3.750
NV	4	+	1.787	1.493	3.705
NV	17	+	1.203	0.877	3.332
AH	4	−	0.869	0.379	2.224
AH	17	−	1.150	0.197	1.143
NV	4	−	2.862	0.864	0.545
NV	17	−	1.359	0.662	1.562

Experiment C					
Subject	N	I	t	r	b
AH	4	+	0.396	0.000	2.747
AH	17	+	0.422	0.000	2.579
NV	4	+	0.725	0.000	1.502
NV	17	+	0.250	0.000	4.356
AH	4	-	0.503	0.000	2.166
AH	17	-	0.489	0.000	2.225
NV	4	-	1.456	0.000	0.748
NV	17	-	0.206	0.000	5.279

Table 5.3.1: Results of Experiment B and C. Some relevant results from experiment A are presented as well. t stands for the threshold (dB); r for the offset (dB) and b for the slope (dB^{-1}).

Comparing the results from experiment A with Experiment B shows us that no significant improvement has been achieved by keeping N, the place of the jump, fixed. Closer examination of the data of Experiment B shows that at a fixed value for N and I great differences in performance occurred by the same subject, varying from 75% correct to chance level. This is probably is caused by the effect that subjects sometimes looses the clue, i.e. suddenly they do not know anymore where to listen. This causes great instability and an overall decrease in performance. Therefore the adaptive method in Experiment C was tried as well.

The results of Experiment C show a much better performance for experiments with fixed N, that is, if we look only at the $I_{70.7}$ data. Since we

assumed $r = 0$ in Experiment C, it is not entirely fair to compare these slopes.

When I< 0, both subjects sometimes reported a *clue reversal*, i.e. the *down*-stimulus was perceived as an *up*-stimulus and vice versa. Closer examination with subject AH showed that by I$=-\infty$ clue-reversal appeared in *log*-complexes at all values of N, whereas in the *lin*-complexes this phenomenon was perceived up to N=10. For N$>$ 10 identification became more difficult (up to 70% correct) and clue reversal disappeared. Instead of clue reversal some kind of timbre recognition was adopted.

5.4 Discussion

Figure 5.3.1 shows that increments in one component of a *log*-complex are much easier perceived than in a *lin*-complex, since the slope of the first one is almost twice as large as the slope of the latter. The offset, however, is in both cases almost the same. (The standard error allows us to let the two offsets coincide). The offset indicated a degree of uncertainty in the signal (Jeffress, 1970, p. 109). In this experiment the uncertainty was due to the fact that the subject did not know a priori the value of N, or so to say, did not know where in the frequency domain the signal would appear. Since the subject knows he's dealing with a rather complicated spectrum, he instanteneously scans along the frequency-axis back and forth, and frequently misses the signal because he's looking at that time at another place (Tanner,1970, p.87). Since this degree of uncertainty is the same for the *lin*- and the *log*-complexes (there are an equal amount of possibilities where the signal can appear), the two offsets coincide.

Figure 5.3.2(a) shows that an increment in the middle of the *log*-complex is more easily perceived than one at either end of the complex. This result was also obtained in previous experiments with analogue multitone complexes (Green *et al.*, 1987; Bernstein & Green, 1987). Profile analysis theory would provide us a possible explanation. In this theory it is assumed that subjects, when listening to a complex spectrum, tend to watch changes in shape, rather than intensity changes of one or more components. Perhaps at edges of the complex profile, analysis is more difficult and performance gets worse.

Another explanation might be, since we assume that subjects scan over

the frequency-axis, that a subject is more concentrated at changes in the center of the complex, thus perceiving changes at the ends less well. Experiments by Green *et al.* (1987) show, on the other hand, that maximum performance always lies near the 1000Hz region, no matter the frequencies of the complex. Sofar there is no real explanation for the shape of the curve in Figure 5.3.2(a). The most we can say is that perhaps subjects are *always* concentrated on frequencies near 1000 Hz.

Figure 5.3.2(b) shows an analogue edge effect for experiments with *lin*-complexes, however less pronounced. One may wonder if this is the same effect as discussed under Figure 5.3.2(a), since two other items become important. First of all, the relative frequency interval $\frac{\Delta f}{f}$ (with $\Delta f = f_2 - f_1$ and $f = (f_1 + f_2)/2$) for small N is rather large. The largest interval occurs for N=1 and is a jump over an octave. Due to the extreme consonance of octaves (Plomp, 1966, Chapter 4), in relation to the total spectrum, identification of the direction of the jump becomes very difficult. Fifth (N=2) and fourth (N=3) jumps suffer as well from this phenomenon, however to a lesser degree. When N becomes larger, the interval becomes smaller and consonance effects become less important. In the second place, near N=10, a small decrease in performance is noticed (see Figure 5.3.2(b) and 5.3.3(b)). In this region $\frac{\Delta f}{f} \approx 0.1$. Furthermore experiments with I=$-\infty$ for *lin*-complexes showed a clue reversal for N$<$ 10 (with a 100% score) and some kind of a timbre recognition for N$>$ 10 (with a 70% score), thus a change appeared near N=10. This change did not appear with the *log*-complexes, since clue reversal was obtained for all N, always with a 100% correct score. The decrease in performance and the change in

the recognition clue probably are due to critical-band effects. However, comparisons of results of these experiments with those of Zwicker *et al.* (1957, Table I) show that effects of critical bands should appear already near N=6. On the other hand, experiments with two simultaneous tones, carried out by Plomp (1966, Chapter 1), shows that the discrimination threshold for two tones with a difference in frequency of 200Hz is near a center frequency of about 2000Hz, which corresponds to our results.

For I< 0 clue reversal was often reported by subjects. This phenomenon can be explained by stating that subjects do not listen to the decremented component in the complex, but rather in some way compares the two complexes, which gives them (1) a jump in the actual direction ($N_1 \rightarrow N_2$), caused by the increment in the components and, directly related, (2) a jump opposite to the actual direction, called the complementary jump, ($N_2 \rightarrow N_1$, see Figure 5.4.1). Since the presence of the complementary jump is stronger than the actual jump (Experiments with random-chord sequences (Chapter 3) have shown that subjects in general do not pay attention to "missing" components) this jump is heard, not the actual jump.

Experiments C (and B as well) clearly show that a decrease of a component in a complex is equally well perceived as an increase.

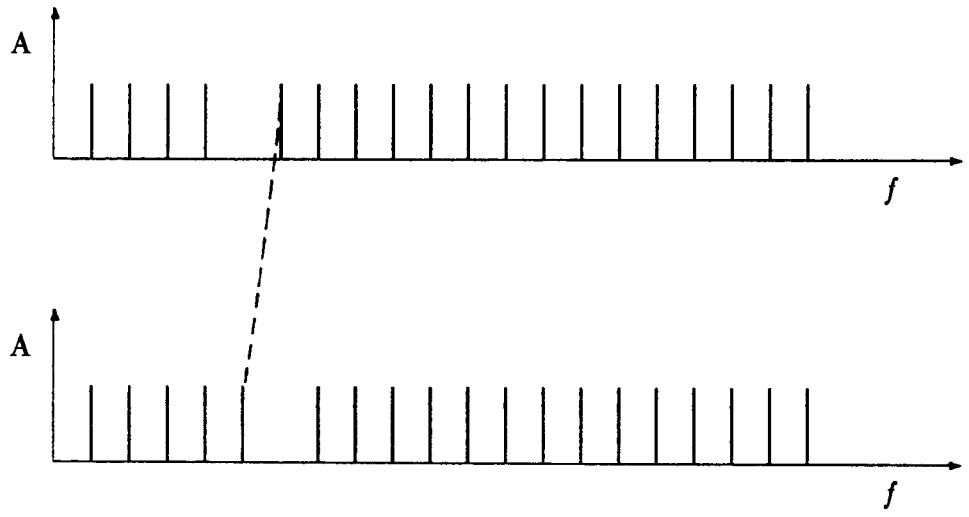


Figure 5.4.1: Representation of an up-stimulus with I= $-\infty$. The upper graph represents the first time frame, the lower graph the second time fame. Instead of perception of the decremented component, a jump is noticed due to the complementary jump (indicated with a dashed line), which results in a clue reversal.

5.5 Conclusions

The results of the experiments allows us to draw the following conclusions:

1. Increments in a component of a *log*-complex are easier perceived than in a *lin*-complex. The identification loss in *lin*-complexes probably is due to consonance-problems for the lower components (N=1 to N=3) and critical band effects for the higher components (N> 10).

2. An increment in the middle of a (*log*-)complex is easier to detect than an increment at one of both ends. It appears that best performance always occurs in the region of 1000Hz. A possible explanation is that subjects, no matter what signal, are focussed on the frequency region near 1 kHz.

3. Uncertainty in the place at the frequency-axis where the signal occurs, causes in both *lin*- and *log*- complexes an offset of about 0.6dB. This agrees with the idea that the size of the offset is related to the number of possible places where the signal might occur (Jeffress, 1970)

4. Performance with *log*-complexes is good, better than in previous analogue experiments of Spiegel *et al.* (1981). Performance is even excellent when N is kept fixed, i.e. when there is no signal uncertainty; much better than in previous experiments of Green *et al.* (1983, 1985, 1987). Since Experiment C was only a preliminary experiment, suggested is to determine the thresholds more carefully, since these values are unbelievebly small.

5. An increment and a decrement in a component of a complex are equally well perceived.

6. Absence or decrements of a component in a complex sometimes causes clue reversals. Apparently one pays less attention to components with smaller amplitudes.

References

Abramowitz, M. & Stegun, I.A. (1956). *Handbook of mathematical functions*, Dover Publications, inc., New York.

Allik, J. & Dzhafarov, E.N. (1984). "Motion direction identification in random cinematograms: a general model", *Journal of Experimental Psychology: Human Perception and Performance 10*, 378-393.

Anstis, S.M. (1970)."Phi movement as subtraction process", *Vision Research 10*, 1411-1430.

Austin, W. (1966). *Music in the 20th Century. From Debussy through Stravinsky*, W.W. Norton, New York.

Beerends, J.G. & Houtsma, A.J.M. (1986). "Pitch identification of simultaneous dichotic two-tone complexes", *Journal of the Acoustical Society of America 80*, 1048-1056.

Bell, H.H. & Lappin, J.S. (1973). "Sufficient conditions for the discrimination of motion", *Perception and Psychophysics 14*, 45-50.

Bernstein, L.R. & Green, D.M. (1987). "Detection of simple and complex changes of spectral shape", *Journal of the Acoustical Society of America 82*, 1587-1592.

de Boer, E. (1976). "On the 'Residue' and auditory pitch perception", in *Handbook of Sensory Physiology, Vol. 5, Part 3*, W.D. Keidel & W.D. Neff eds., Springer, New York.

Bregman, A.S. & Campbell, J. (1971)."Primary auditory stream segregation and perception of order in rapid sequence of tones", *Journal of Experimental Psychology 89*, 244-249.

Burns, E.M. & Ward, W.D. (1978). "Categorical perception- Phenomenon

or epiphenomenon: evidence from experiments in the perception of melodic musical intervals", *Journal of the Acoustical Society of America 63*, 456-468.

Butler, D. (1979). "A further study of melodic channeling", *Perception & Psychophysics 25*, 264-268.

Deutsch, D. (1975). "Two-channel listening to musical scales", *Journal of the Acoustical Society of America 57*, 1156-1160.

Deutsch, D. (1980). "The processing of structured and unstructured tonal sequences", *Perception & Psychophysics 28*, 381-389.

Deutsch, D. (1982). "Grouping Mechanisms in Music", in *The Psychology of Music, Chapter 4*, Edited by D. Deutsch, Academic Press, New York.

Doehring, D.G. (1971). "Discrimination of simultaneous and successive pure tones by musical and nonmusical subjects", *Psychonomic Science 22*, 209-210.

van Doorn A.J. & Koenderink, J.J. (1982). "Spatial properties of the visual detectability of moving spatial white noise", *Experimental Brain Research 45*, 189-195.

van Doorn, A.J. & Koenderink, J.J. (1984). "Spatiotemporal integration in the detection of coherent motion", *Vision Research 24*, 47-54.

Doughty, J.M. & Garner, W.M. (1948). "Pitch characteristics of short tones II: pitch as a function of duration", *Journal of Experimental Psychology 38*, 478-494.

Dowling, W.J. (1973). "The perception of interleaved melodies", *Cognitive Psychology 5*, 322-337.

Eiting, M. (1984). *Aspects of the cognition of tonal music*, Doctoral Dissertation, Univ. of Amsterdam.

Goldstein, J.L. (1967). "Auditory nonlinearity", *Journal of the Acoustical Society of America 41*, 676-689.

Green, D.M. & Swets, J.A. (1966). *Signal detection theory and psychophysics*, Krieger, New York.

Green, D.M.; Kidd Jr., G. & Picardi, M.C. (1983). "Successive versus simultaneous comparison in auditory intensity discrimination", *Journal of the Acoustical Society of America 73*, 639-643.

Green, D.M. & Kidd Jr., G. (1983). "Further studies of auditory profile analysis", *Journal of the Acoustical Society of America 73*, 1260-1265.

Green, D.M. & Mason, C.R. (1985). "Auditory profile analysis: Frequency, phase, and Weber's Law", *Journal of the Acoustical Society of America 77*, 1155-1161.

Green, D.M.; Onsan, Z.A. & Forrest, T.G. (1987). "Frequency effects in profile analysis and detecting complex spectral changes", *Journal of the Acoustical Society of America 81*, 692-699.

Hartmann, W.M. (1978). "The effect of amplitude envelope on the pitch of sinewave tones", *Journal of the Acoustical Society of America 63*, 1105-1113.

Jeffress, L.A. (1970). "Masking", in *Foundations of Modern Auditory Theory, Vol. 1, Chapter 3*, Edited by J.V. Tobias, Academic Press, New York.

Julesz, B. & Bosch, C. (1966). *Studies on Visual Texture and Binocular Depth Perception.* A computer-generated movie series containing monocular and binocular movies. Bell Telephone Laboratories, Inc.

Julesz, B. (1971). *Foundation of Cyclopean Perception*, Chicago University Press, Chicago.

Kimura, D. (1964). "Left-right differences in the perception of melodies", *Quarterly Journal of Experimental Psychology 16*, 355-358.

Korte, A. (1915). "Kinematoskopische Untersuchungen", *Zeitschrift für Psychologie 72*, 193-296.

Larkin, W.D. (1978). "Pitch shifts following tone adaptation", *Acustica 41*, 110-116.

Levitt, H. (1971). "Transformed up-down methods in psychoacoustics", *Journal of the Acoustical Society of America 49*, 467-477.

Levitt, H. & Rabiner, L.R. (1967). "Use of a sequential strategy in intelligibility testing", *Journal of the Acoustical Society of America 42*, 609-612.

Lippmann, R.P.; Braida, L.D. & Durlach, N.I. (1976). "Intensity perception. V. Effect of payoff matrix on absolute identification", *Journal of the Acoustical Society of America 59*, 129-134.

Nakayama, K, & Silverman, G. "Temporal and spatial characteristics of the upper displacement limit for motion in random dots", *Vision Research 24*, 293-300.

van Noorden, L.P.A.S. (1975). *Temporal Coherence in the Perception of Tone Sequences*, Doctoral Dissertation, Eindhoven University of Technology.

Plomp, R. (1966). *Experiments on Tone Perception*, Doctoral Dissertation, Utrecht University.

Risset, J-C. (1986). "Pitch and rhythm paradoxes: Comment on 'Auditory paradox based on fractal waveform' [J. Acoust. Soc. Am. 79, 186-189, 1986]", *Journal of the Acoustical Society of America 80*, 961-962.

Rossing, T.D. & Houtsma, A.J.M. (1986). "Effects of signal envelope on

the pitch of short sinusoidal tones", *Journal of the Acoustical Society of America 79*, 1926-1933.

Scharf, B. & Houtsma, A.J.M. (1986). "Audition II: Loudness, Pitch, Localization, Aural Distortion, Pathology", in *Handbook of Perception and Human Performance, Vol. 1*, K.R. Boff, L. Kaufman & J.P. Thomas eds., J. Wiley, New York.

Schroeder, M.R. (1986). "Auditory paradox based on fractal waveform", Journal of the Acoustical Society of America 79, 186-189.

Shelton, B.R.; Picardi, M.C. & Green, D.M. (1982). "Comparison of three adaptive psychophysical procedures", *Journal of the Acoustical Society of America 71*, 1527-1533.

Shepard, R.N. (1964). "Circularity in the judgments of relative pitch", *Journal of the Acoustical Society of America 36*, 2346-2353.

Spiegel, M.F.; Picardi, M.C. & Green, D.M. (1981). "Signal and masker uncertainty in intensity discrimination", *Journal of the Acoustical Society of America 70*, 1015-1019.

Spiegel, M.F. & Green, D.M. (1982). "Signal and masker uncertainty with noise maskers of varying duration, bandwith, and center frequency", *Journal of the Acoustical Society of America 71*, 1204-1210.

Stevens, S.S. (1935). "The relation of pitch to intensity", *Journal of the Acoustical Society of America 6*, 150-154.

Stevens, S.S. & Volkmann, J. (1940). "The relation of pitch to frequency", *American Journal of Psychology 53*, 329-353.

Tanner Jr., W.P. & Sorkin, R.D. (1970). "The theory of signal detectability" in *Foundations of Modern Auditory Theory, Vol. 2, Chapter 2,* Edited

by J.V. Tobias, Academic Press, New York.

Terhardt, E. & Fastl, H. (1971). "Zum Einfluss von Störtönen und Störgeräuschen auf die Tonhöhe von Sinustönen", *Acustica 25*, 53-61.

Verschuure, J. & van Meeteren, A.A. (1975). "The effect of intensity on pitch", *Acustica 32*, 33-44.

Wertheimer, M. (1912). "Experimentelle Studien über das Sehen von Bewegung", *Zeitschrift für Psychologie 61*, 161-265.

Zwicker, E. (1955). "Der ungewöhnliche Amplitudengang der nichtlinearen Verzerrungen des Ohres", *Acustica 5*, 67-74.

Zwicker, E.; Flottorp, G. & Stevens, S.S. (1957). "Critical Band Width in Loudness Summation", *Journal of the Acoustical Society of America 29*, 548-557.

Acknowledgements

First of all I am greatfully indebted to my coach Aad Houtsma for his encouraging support, useful advices and help, which resulted in writing this report.

Furthermore Berry Eggen, Jacques Grupa, Jim Juola, Rick de Lange and Jacek Smurzynski are acknowledged for their assistance as subjects.

Finally I am indebted to Theo de Jong for his invaluable technical assistance.

Niek Versfeld

I Appendix I

Appendix I shows the calculation of E[TSC] (Appendix Ia), Var[TSC] (Appendix Ib) and Cov[TSC] (Appendix Ic).

I.1 Appendix Ia

This appendix shows the computation of E[TSC]. According to general rules:

$$\mathrm{E[TSC]} := \mathrm{E}[\sum_{\mathrm{d}} \mathrm{c(d)}] = \sum_{\mathrm{d}} \mathrm{E[c(d)]}. \tag{1}$$

Breaking down $E[c(d)]$ into its partial contributions corresponding to the four different dipole forms h with the associated probabilities Pd_h, we can write Equation (1) as:

$$\mathrm{E[TSC]} = \sum_{\mathrm{d}} \mathrm{E[c(d)]} = \sum_{\mathrm{d}} \sum_{\mathrm{h}} \mathrm{E[c(d_h)]Pd_h}. \tag{2}$$

Since only dipoles with form h_{11} have a non-zero contribution, Equation (2) can be further simplified to:

$$\sum_{\mathrm{d}} \sum_{\mathrm{h}} \mathrm{E[c(d_h)]Pd_h} = \sum_{\mathrm{d}} \mathrm{E[c(d_{11})]Pd_{11}}. \tag{3}$$

According to the symmetry principle, contributions of any two symmetrical dipoles are identical but differ in sign :

$$\mathrm{E[c(d_{11}^-)]} = -\mathrm{E[c(d_{11}^+)]}, \tag{4}$$

where d_{11}^- and d_{11}^+ have displacement vectors $(-f,t)$ and (f,t), respectively, and are both of the h_{11} form. We now split up the sum of Equation (3) into two parts, one being the sum of positive $((f,t))$, the other the sum of negative $((-f,t))$ dipoles. We finally can rewrite the summation over only

positive dipoles since the total number of positive dipoles equals the total number of negative dipoles in each RCS:

$$\begin{aligned}\sum_{d} \mathrm{E}[\mathrm{c}(\mathrm{d}_{11})]\mathrm{Pd}_{11} &= \sum_{d>0} \mathrm{E}[\mathrm{c}(\mathrm{d}^{+}_{11})]\mathrm{Pd}^{+}_{11} + \sum_{d<0} \mathrm{E}[\mathrm{c}(\mathrm{d}^{-}_{11})]\mathrm{Pd}^{-}_{11} \\ &= \sum_{d>0} \mathrm{E}[\mathrm{c}(\mathrm{d}^{+}_{11})](\mathrm{Pd}^{+}_{11} - \mathrm{Pd}^{-}_{11}). \qquad (5)\end{aligned}$$

Since only the shortest dipoles contribute, we simply need to sum over all $(1,1)$-dipoles. According to the homogeneity principle, all these contributions are equal, so that :

$$\mathrm{E[TSC]} = \mathrm{ME}^{+}_{11}(\mathrm{Pd}^{*+}_{11} - \mathrm{Pd}^{*-}_{11}), \qquad (6)$$

where d^{*+}_{11} is a dipole with displacement vector $(1,1)$ and form h_{11}, d^{*-}_{11} a dipole with displacement vector $(-1,1)$ and form h_{11}, and M the total number of $(1,1)$-dipoles in an RCS. E^{+}_{11} is a short notation for $E[c(d^{+}_{11})]$. The chance that a $(1,1)$-dipole is in h_{11} form (Pd^{*+}_{11}) is either $\frac{1}{2}P$ or $\frac{1}{4}$, depending on the motion direction N ; Pd^{*-}_{11} then is $\frac{1}{4}$ or $\frac{1}{2}P$, respectively. Assuming that $N = 1$, we get for E[TSC] :

$$\mathrm{E[TSC]} = \mathrm{ME}^{+}_{11}(\frac{1}{2}\mathrm{P} - \frac{1}{4}). \qquad (7)$$

This equation corresponds to Equation (4) in the text. For N=-1, the same expression is obtained for E[TSC] except for an opposite sign.

I.2 Appendix Ib

This appendix shows the calculation of Var[TSC]. According to general rules:

$$\begin{aligned} \mathrm{Var[TSC]} &:= \mathrm{Var}[\sum_{d} c(d)] \\ &= \sum_{d} \mathrm{Var}[c(d)] + 2 \sum_{i} \sum_{j \neq i} \mathrm{Cov}[c_i(d), c_j(d)] \end{aligned} \tag{8}$$

and

$$\mathrm{Var}[c(d)] = \mathrm{E}[c^2(d)] - \mathrm{E}^2[c(d)], \tag{9}$$

where

$$\begin{aligned} \mathrm{E}[c^2(d)] &= \sum_{h} \mathrm{E}[c^2(d_h)] \mathrm{P}d_h \\ \mathrm{E}^2[c(d)] &= \{\sum_{h} \mathrm{E}[c(d_h)] \mathrm{P}d_h\}^2. \end{aligned}$$

Pd_h is the probability that the dipole is of form h, and the summation is done over all four possible forms h. Using the fact that only $E[c(d_{11})]$ is non-zero, we can write Equation (9) as:

$$\mathrm{Var}[c(d)] = \mathrm{E}^2[c(d_{11})] \mathrm{P}d_{11}(1 - \mathrm{P}d_{11}) + \sum_{h} \mathrm{Var}[c(d_h)] \mathrm{P}d_h. \tag{10}$$

With the aid of the symmetry principle we now write the sum of variances over only the positive dipoles. If, in addition, we only consider the shortest (1,1) dipoles, we obtain:

$$\begin{aligned} \sum_{d} \mathrm{Var}[c(d)] &= \sum_{d^*>0} (\sum_{h} \{\mathrm{Var}[c(d_h^{*+})](\mathrm{P}d_h^{*+} + \mathrm{P}d_h^{*-})\} \\ &\quad + \mathrm{E}^2[c(d_{11}^{*+})](\mathrm{P}d_{11}^{*+}(1 - \mathrm{P}d_{11}^{*+}) + \mathrm{P}d_{11}^{*-}(1 - \mathrm{P}d_{11}^{*-})\} \end{aligned} \tag{11}$$

where the notations d^{*+} and d^{*-} designate dipoles with displacement vectors of (1,1) and (-1,1), respectively, as defined in Appendix Ia. Regardless of whether N=1 (*i.e.* $Pd_{11}^{*+} = \frac{1}{2}P$ and $Pd_{11}^{*-} = \frac{1}{4}$) or N= -1 ($Pd_{11}^{*+} = \frac{1}{4}$ and $Pd_{11}^{*-} = \frac{1}{2}P$), we obtain the final result:

$$\sum_{d} \mathrm{Var}[c(d)] = \frac{1}{4} M (E_{11}^{+})^2 \{(2P+1)V_j + (2Q+1)V_{nj} + P + PQ + \frac{3}{4}\}, \quad (12)$$

where M is the number of dipoles with displacement vectors (1,1) in an RCS, E_{11}^{+} is a short notation for $E[c(d_{11}^{++})]$, Q=1-P, and

$$V_j := \frac{\mathrm{Var}[c(d_{11}^{+})] + \mathrm{Var}[c(d_{00}^{+})]}{(E_{11}^{+})^2} \quad (13)$$

$$V_{nj} := \frac{\mathrm{Var}[c(d_{10}^{+})] + \mathrm{Var}[c(d_{01}^{+})]}{(E_{11}^{+})^2}. \quad (14)$$

In order to calculate Var[TSC] we need to find an expression for the covariance. This will be done in Appendix Ic. Var[TSC] then is the sum of Equation (12) from Appendix Ib and two times Equation (52) from Appendic Ic.

I.3 Appendix Ic

This appendix contains the computation of Cov[TSC]. According to general rules:

$$\begin{aligned} \mathrm{Cov}[TSC] &:= \sum_{i} \sum_{j \neq i} \mathrm{Cov}[c_i(d), c_j(d)] \\ &= \sum_{i} \sum_{j \neq i} \{E[c_i(d) \cdot c_j(d)] - E[c_i(d)] \cdot E[c_j(d)]\}. \end{aligned} \quad (15)$$

Again, breaking up E[c(d)] into partial contributions associated with the four different dipole forms, we obtain:

$$\begin{aligned} \mathrm{Cov}[c_i(d), c_j(d)] &= \sum_{h_1} \sum_{h_2} \{E[c_i(d_{h_1}) c_j(d_{h_2})] P_{ij}d \\ &\quad - E[c_i(d_{h_1})] E[c_i(d_{h_2})] P_i d P_j d\}, \end{aligned} \quad (16)$$

where $P_{ij}d$ means the joint probability that the ith dipole is in one form h_1 *and* the jth dipole in some other form h_2. $P_i d$ means that dipole i is in form h_1. According to the independence principle:

$$E[c_i(d_{h_1}) \cdot c_j(d_{h_2})] = E[c_i(d_{h_1})]E[c_j(d_{h_2})], \tag{17}$$

so that Equation (16) becomes:

$$Cov[c_i(d), c_j(d)] = \sum_{h_1} \sum_{h_2} E[c_i(d_{h_1})]E[c_j(d_{h_2})](P_{ij}d - P_i d P_j d). \tag{18}$$

We now use the fact that only $E[c(d_{11})]$ is non-zero, which gives us:

$$Cov[c_i(d), c_j(d)] = E[c_i(d_{11})]E[c_j(d_{11})](P_{ij}d_{11} - P_i d_{11} P_j d_{11}). \tag{19}$$

As in the previous appendices, we now restrict ourselves to summation over only the (1,1)- and (-1,1)-dipoles, which yields four different ways in which the two dipoles can be grouped, namely (1) both dipoles are (1,1)-dipoles, (2) one (1,1)-dipole is grouped with one (-1,1)-dipole, (3) one (-1,1)-dipole is grouped with one (1,1)-dipole, and (4) both dipoles are (-1,1)-dipoles. So we may write Cov[TSC] as:

$$\begin{aligned} Cov[TSC] \;=\; & \sum_{i^+} \sum_{j^+ \neq i^+} E_{11}^+ E_{11}^+ (P_{ij}^{++} d_{11}^{**} - P_i d_{11}^{*+} P_j d_{11}^{*+}) \\ & + \sum_{i^+} \sum_{j^-} E_{11}^+ E_{11}^- (P_{ij}^{+-} d_{11}^{**} - P_i d_{11}^{*+} P_j d_{11}^{*-}) \\ & + \sum_{i^-} \sum_{j^+} E_{11}^- E_{11}^+ (P_{ij}^{-+} d_{11}^{**} - P_i d_{11}^{*-} P_j d_{11}^{*+}) \\ & + \sum_{i^-} \sum_{j^- \neq i^-} E_{11}^- E_{11}^- (P_{ij}^{--} d_{11}^{**} - P_i d_{11}^{*-} P_j d_{11}^{*-}), \end{aligned} \tag{20}$$

where i^+ and j^+ apply to all positive shortest-dipole pairs with form h_{11}, i^- and j^- to the negative ones. $P_{ij}^{+-} d_{11}^{**}$, for instance, is the joint probability that the ith dipole is of form h_{11} with displacement vector (1,1) and the jth dipole is of form h_{11} with displacement vector (-1,1). E_{11}^+andE_{11}^- are the short notations for $E[c_i(d_{11}^+)]$ and $E[c_i(d_{11}^-)]$, respectively. We observe that $E_{11}^- = -E_{11}^+$ and that $P_{ij}^{+-} d_{11}^{**} = P_{ij}^{-+} d_{11}^{**}$.

What we obviously have to do is to search for those dipole pairs where $P_{ij} d_{11}^{**} \neq P_i d_{11}^* P_j d_{11}^*$. This appears to be the case only when the two dipoles

can be connected by a so-called *chain* (shown in Figure Ic.1. A chain is a diagonal in an RCS into the direction of motion (N). Two elements (sf, st) lying apart from each other on a chain have the same states with probability P_s:

$$P_s = \sum_{2i \leq s}^{s} \begin{pmatrix} s \\ 2i \end{pmatrix} P^{s-2i} Q^{2i}. \tag{21}$$

So for:

$$\begin{aligned} s = 0: &\quad P_1 = 1 \\ s = 1: &\quad P_2 = P(\equiv SRP) \\ s = 3: &\quad P_3 = P^2 + Q^2 \\ s = 4: &\quad P_4 = P^3 + 3PQ^2, ect. \end{aligned} \tag{22}$$

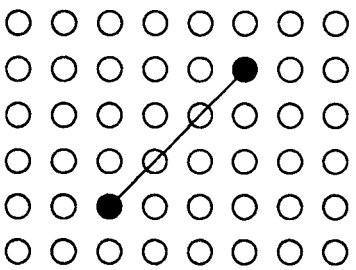


Figure Ic.1: A chain in a RCS. The chainlength s=3.

The four terms in Equation (20) give rise to eight different ways in which two dipoles can be connected by a chain: The *aligned* version results from the first term of Equation (20), the γ_1 and γ_2 version from the second term, the γ_3 and γ_4 version from the third and the *parallellogram*, z_1 and z_2 version from the fourth. The eight versions are discussed next, where for simplicity we take N=+1.

1. The *aligned* version is shown in Figure Ic.2.

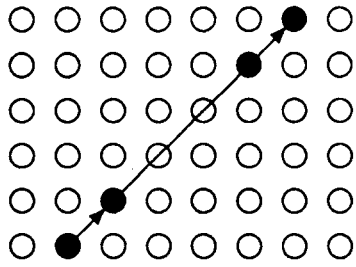


Figure Ic.2: Structure of the Aligned version

We can see that $P_{ij}^{++}d_{11}^{**}$ (the probability that *both* dipoles have form h_{11}) has the value $\frac{1}{2}P \cdot P_s \cdot P = \frac{1}{2}P^2P_s$. We also notice that $P_i d_{11}^{*+} = P_j d_{11}^{*+} = \frac{1}{2}P$, so the covariance term for the aligned version will be (obtained by the substitution in Eq. C6):

$$\mathrm{Cov}^{(\mathrm{ali})} = \frac{1}{4}\mathrm{P}^2\mathrm{E}_{11}^{+}\mathrm{E}_{11}^{+}\sum_{\mathrm{s}}\mathrm{N}_{\mathrm{s}}^{(\mathrm{ali})}(2\mathrm{P}_{\mathrm{s}}-1), \tag{23}$$

where $N_s^{(ali)}$ is the frequency of occurrence of an aligned dipole-pair with a connecting chain of length s in a RCS. $N_s^{(ali)}$ depends on the size of an RCS (number of frames and elements) and on the way the RCS is constructed (i.e. circular or non-circular), because this determines values of s in the sum.

2. The γ_1-version is shown in Figure Ic.3.

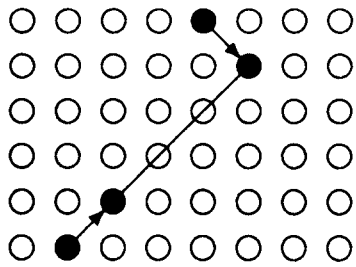


Figure Ic.3: Structure of the γ_1-version

As can be seen:

$$P_{ij}^{+-} d_{11}^{**} = \frac{1}{2} P \cdot P_s \cdot \frac{1}{2} \tag{24}$$

$$P_i d_{11}^{*+} = \frac{1}{2} P \tag{25}$$

$$P_i d_{11}^{*-} = \frac{1}{4} \tag{26}$$

so:

$$Cov^{(\gamma_1)} = \frac{1}{8} P E_{11}^{+} E_{11}^{-} \sum_s N_s^{(\gamma_1)} (2P_s - 1) \tag{27}$$

3. The γ_2-version is shown in Figure Ic.4.

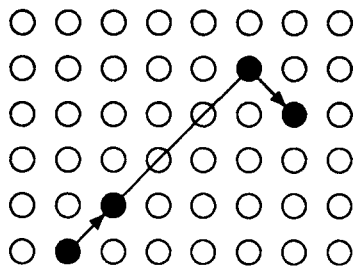


Figure Ic.4: Structure of the γ_2-version

As can be seen:

$$P_{ij}^{+-} d_{11}^{**} = \frac{1}{2} P \cdot P_s \cdot \frac{1}{2} \tag{28}$$

$$P_i d_{11}^{*+} = \frac{1}{2} P \tag{29}$$

$$P_j d_{11}^{*-} = \frac{1}{4} \tag{30}$$

so:

$$Cov^{(\gamma_2)} = \frac{1}{8} P E_{11}^{+} E_{11}^{-} \sum_s N_s^{(\gamma_2)} (2P_s - 1) \tag{31}$$

4. The γ_3-version is shown in Figure Ic.5.

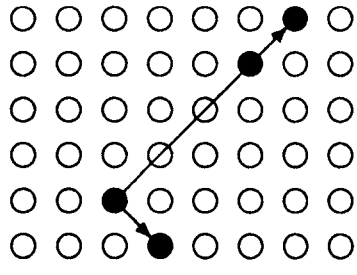


Figure Ic.5: Structure of the γ_3-version

As can be seen:

$$P_{ij}^{-+} d_{11}^{**} = \frac{1}{2} P \cdot P_s \cdot \frac{1}{2} \tag{32}$$

$$P_i d_{11}^{*-} = \frac{1}{4} \tag{33}$$

$$P_j d_{11}^{*+} = \frac{1}{2} P \tag{34}$$

so:

$$Cov^{(\gamma_3)} = \frac{1}{8} P E_{11}^{-} E_{11}^{+} \sum_s N_s^{(\gamma_3)} (2P_s - 1) \tag{35}$$

5. The γ_4-version is shown in Figure Ic.6.

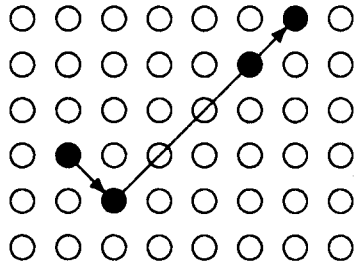


Figure Ic.6: Structure of the γ_4-version

As can be seen:

$$P_{ij}^{-+}d_{11}^{**} = \frac{1}{2}P \cdot P_s \cdot \frac{1}{2} \tag{36}$$

$$P_i d_{11}^{*-} = \frac{1}{4} \tag{37}$$

$$P_j d_{11}^{*+} = \frac{1}{2}P \tag{38}$$

so:

$$Cov^{(\gamma_4)} = \frac{1}{8}PE_{11}^{-}E_{11}^{+}\sum_s N_s^{(\gamma_4)}(2P_s - 1) \tag{39}$$

6. The Z_1-version is shown in Figure Ic.7.

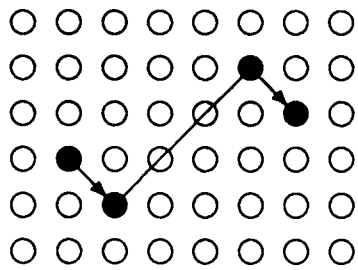


Figure Ic.7: Structure of the Z_1-version

As can be seen:

$$P_{ij}^{--}d_{11}^{**} = \frac{1}{2} \cdot \frac{1}{2}P_s \cdot \frac{1}{2} \tag{40}$$

$$P_i d_{11}^{*-} = \frac{1}{4} \tag{41}$$

$$P_j d_{11}^{*-} = \frac{1}{4} \tag{42}$$

so:

$$Cov^{(Z_1)} = \frac{1}{16} E_{11}^{-} E_{11}^{-} \sum_s N_s^{(Z_1)} (2P_s - 1) \tag{43}$$

7. The Z_2-version is shown in Figure Ic.8.

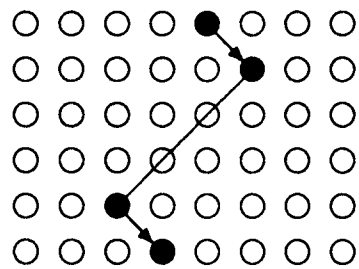


Figure Ic.8: Structure of the Z_2-version

As can be seen:

$$P_{ij}^{--} d_{11}^{**} = \frac{1}{2} \cdot \frac{1}{2} P_s \cdot \frac{1}{2} \tag{44}$$

$$P_i d_{11}^{*-} = \frac{1}{4} \tag{45}$$

$$P_j d_{11}^{*-} = \frac{1}{4} \tag{46}$$

so:

$$Cov^{(Z_2)} = \frac{1}{16} E_{11}^{-} E_{11}^{-} \sum_s N_s^{(Z_2)} (2P_s - 1) \tag{47}$$

8. The *parallellogram* version is shown in figure Ic.9.

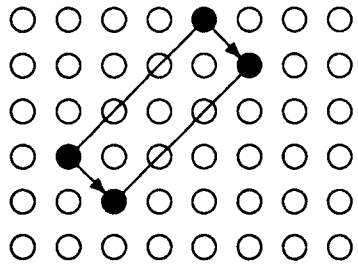


Figure Ic.9: Structure of the parallellogram-version

As can be seen:

$$P_{ij}^{--}d_{11}^{**} = \frac{1}{2}P_s \cdot \frac{1}{2}P_s \tag{48}$$

$$P_i d_{11}^{*-} = \frac{1}{4} \tag{49}$$

$$P_j d_{11}^{*-} = \frac{1}{4} \tag{50}$$

so:

$$Cov^{(par)} = \frac{1}{16}E_{11}^{-}E_{11}^{-}\sum_s N_s^{(par)}(4P_s^2 - 1) \tag{51}$$

We observe that all except the γ versions have a positive value (since $E_{11}^{-} \cdot E_{11}^{+}$ is negative). All eight sub-covariances now sum to the total covariance term:

$$Cov[TSC] = \sum_{k=1}^{8} Cov^{(k)} \tag{52}$$

The variance of the Total Sum of Contribution is the sum of Equation (12) of Appendic Ib and two times Equation (52). We see that Cov[TSC] depends only on the SRP, the number of frames and elements in a RCS and on its construction (circular or non-circular).

II Appendix II

In this appendix the data of the individual subjects are presented as obtained from the experiments described in Chapter 4.

The results of each subject are presented over two horizontal lines. The first line (denoted with a "g") is the number of correct responses. The second line (denoted with an "N") is the number of trials for that condition. Furthermore the results are devided in *up*-stimuli (denoted with an "u") and *down*-stimuli (denoted with an "d"). The last two rows are the pooled data.

Experiment A Score versus Pn

10		9		8		7		6		5		4		3		2		1		0		Pn	
U	d	U	d	U	d	U	d	U	d	U	d	U	d	U	d	U	d	U	d	U	d		
43	48	119	92	152	168	168	193	160	235	184	249	273	216	157	115	106	168	60	60	60	60	g	AH
84	108	216	180	268	248	264	252	232	284	232	284	296	220	160	116	108	168	60	60	60	60	N	
30	60	95	130	100	211	129	211	144	243	168	261	249	212	152	113	108	166	60	60	60	60	g	JG
84	108	216	180	268	248	264	252	232	284	232	284	296	220	160	116	108	168	60	60	60	60	N	
26	68	116	110	166	145	194	168	181	220	195	230	267	206	148	115	106	163	58	59	60	60	g	RL
84	108	216	180	268	248	264	252	232	284	232	284	296	220	160	116	108	168	60	60	60	60	N	
53	45	127	93	166	168	194	194	161	238	189	256	270	207	160	115	108	168	60	60	60	60	g	NV
84	108	216	180	268	248	264	252	232	284	232	284	296	220	160	116	108	168	60	60	60	60	N	
152	221	457	425	584	692	685	766	646	936	736	9961057	841	617	458	428	665	238	239	240	240		g	TOT
336	432	864	7201072	99210561008		9281136		92811361184		880	640	464	432	672	240	240	240	240		N		TOT	

****

Experiment A Score versus N

6		5		4		3		2		1		0		N	
U	D	U	D	U	D	U	D	U	D	U	D	U	D		
135	149	198	212	266	290	269	321	249	284	227	192	138	156	G	AH
304	300	348	320	368	368	324	348	260	292	236	196	140	156	N	
94	194	152	257	215	322	241	322	233	281	224	195	136	156	G	JG
304	300	348	320	368	368	324	348	260	292	236	196	140	156	N	
129	174	221	191	290	267	271	296	243	272	223	190	138	154	G	RL
304	300	348	320	368	368	324	348	260	292	236	196	140	156	N	
160	152	220	195	276	292	275	326	253	288	224	195	140	156	G	NV
304	300	348	320	368	368	324	348	260	292	236	196	140	156	N	
518	669	791	855104711711056126		5	9781125		898	772	552	622			G	TOT
121612001392128014721472129613921040						1168		944	784	560	624			N	TOT

****

Experiment A Score versus D

5		4		3		2		1		0		D	
U	D	U	D	U	D	U	D	U	D	U	D		
135	149	87	68	181	205	236	211	268	380	575	591	G	AH
304	300	120	80	316	308	292	252	336	436	612	604	N	
94	194	72	73	148	250	189	229	247	393	545	588	G	JG
304	300	120	80	316	308	292	252	336	436	612	604	N	
129	174	84	70	214	184	238	188	275	358	575	570	G	RL
304	300	120	80	316	308	292	252	336	436	612	604	N	
160	152	94	71	198	189	245	205	268	392	583	595	G	NV
304	300	120	80	316	308	292	252	336	436	612	604	N	
518	669	337	282	741	828	908	8331058152322782344					G	TOT
12161200		480	320126412321168100813441744244824							16		N	TOT

****

Experiment A score versus Sum(D)

-4		-3		-2		-1		0		1		2		3		4		5		6		7		8		sum(D
U	d	U	d	U	d	U	d	U	d	U	d	U	d	U	d	U	d	U	d	U	d	U	d	U	d	
2	0	0	4	8	19	80	75	293	305	236	305	260	284	230	239	158	194	159	139	40	32	12	8	4	0	g AH
8	0	0	4	16	28	152	124	512	496	356	376	296	324	256	248	168	200	160	140	40	32	12	8	4	0	N
3	0	0	4	6	26	64	101	226	377	229	338	208	274	204	235	148	193	154	139	37	32	12	8	4	0	g JG
8	0	0	4	16	28	152	124	512	496	356	376	296	324	256	248	168	200	160	140	40	32	12	8	4	0	N
6	0	0	4	12	20	97	75	297	321	256	288	259	251	223	227	158	182	153	136	38	32	12	8	4	0	g RL
8	0	0	4	16	28	152	124	512	496	356	376	296	324	256	248	168	200	160	140	40	32	12	8	4	0	N
7	0	0	4	10	17	81	90	329	294	251	313	260	278	238	238	156	190	160	140	40	32	12	8	4	0	g NV
8	0	0	4	16	28	152	124	512	496	356	376	296	324	256	248	168	200	160	140	40	32	12	8	4	0	N
18	0	0	16	36	82	322	34111451297			9721244		9871087		895	939	620	759	626	554	155	128	48	32	16	0	g TOT
32	0	0	16	64	112	608	4962048198414241504118412961024							992	672	800	640	560	160	128	48	32	16	0	N	TOT

****

Experiment B Score versus N

6		5		4		3		2		1		0		N	
U	D	U	D	U	D	U	D	U	D	U	D	U	D		
2	12	50	68	175	180	239	237	0	0	0	0	0	0	G	AH
12	12	72	72	180	180	240	240	0	0	0	0	0	0	N	
0	10	51	69	168	180	240	237	0	0	0	0	0	0	G	JG
12	12	72	72	180	180	240	240	0	0	0	0	0	0	N	
3	7	58	68	174	176	240	238	0	0	0	0	0	0	G	RL
12	12	72	72	180	180	240	240	0	0	0	0	0	0	N	
4	6	61	70	178	178	240	239	0	0	0	0	0	0	G	NV
12	12	72	72	180	180	240	240	0	0	0	0	0	0	N	
9	35	220	275	695	714	959	951	0	0	0	0	0	0	G	TOT
48	48	288	288	720	720	960	960	0	0	0	0	0	0	N	TOT

****

EXPERIMENT B SCORE VERSUS D

5		4		3		2		1		0	D	
U	D	U	D	U	D	U	D	U	D	U	D	
2	12	18	23	67	81	153	156	179	178	47	47	G AH
12	12	24	24	84	84	156	156	180	180	48	48	N
0	10	21	24	66	81	149	154	175	179	48	48	G JG
12	12	24	24	84	84	156	156	180	180	48	48	N
3	7	21	23	73	81	151	153	179	178	48	47	G RL
12	12	24	24	84	84	156	156	180	180	48	48	N
4	6	24	24	73	82	154	154	180	179	48	48	G NV
12	12	24	24	84	84	156	156	180	180	48	48	N
9	35	84	94	279	325	607	617	713	714	191	190	G TOT
48	48	96	96	336	336	624	624	720	720	192	192	N

****

III Appendix III

In this appendix the data of the individual subjects are presented as obtained from the experiments described in Chapter 5.

The results for each subject for either the *lin*-complex or *log*-complex are presented separately. Each time the first 19 lines present the number of trials for one value of N with correct response. The line denoted with "N per component" gives the number of trials per component. Furthermore the results are devided in *up*-stimuli (denoted with an "u") and *down*-stimuli (denoted with an "d"). The last two rows are the pooled data for all components.

1.0		1.5		2.0		2.5		3.0		4.0		dB			
u	d	u	d	u	d	u	d	u	d	u	d	u:up	d:down		
15	22	18	27	23	22	0	0	24	33	11	15	Subject AH			
16	19	9	23	20	28	0	0	22	34	15	15				
25	23	26	27	22	33	0	0	33	35	14	14	log-complex			
22	20	21	26	31	33	0	0	33	35	15	15				
19	19	29	27	30	34	0	0	33	35	15	15				
23	31	28	30	29	34	0	0	34	35	14	15				
22	25	27	29	32	32	0	0	32	35	13	15				
23	22	28	30	35	33	0	0	33	35	15	15				
28	22	28	27	34	33	0	0	34	35	15	15				
24	26	29	30	34	34	0	0	33	35	14	15				
13	22	24	28	29	32	0	0	35	35	15	15				
22	30	26	35	28	35	0	0	35	35	15	15				
29	28	35	34	35	35	0	0	35	35	15	15				
19	20	30	30	35	34	0	0	35	34	15	15				
26	31	33	34	33	35	0	0	34	35	14	15				
24	21	34	29	34	34	0	0	34	35	15	14				
23	20	32	26	33	30	0	0	33	34	14	14				
18	18	26	24	28	21	0	0	34	28	15	14				
24	18	27	21	28	17	0	0	32	21	15	15	N correct	per	component	
35	35	35	35	35	35	0	0	35	35	15	15	N tot	per	component	
415	437	510	537	573	589	0	0	618	639	274	281	N correct			
665	665	665	665	665	665	0	0	665	665	285	285	N tot			
**********															
11	19	27	32	26	31	0	0	26	29	27	27	Subject RL			
18	25	23	18	31	29	0	0	27	24	29	28				
17	30	24	27	30	28	0	0	28	28	29	30	log-complex			
19	18	25	30	29	33	0	0	30	28	30	28				
20	25	29	32	34	30	0	0	29	29	28	29				
16	19	29	29	31	33	0	0	30	29	30	28				
18	25	27	32	32	32	0	0	27	27	29	28				
21	24	25	35	32	33	0	0	29	28	29	30				
29	25	29	35	32	34	0	0	30	29	30	29				
22	26	26	29	33	34	0	0	29	29	30	28				
23	19	29	35	35	31	0	0	29	29	28	29				
15	24	29	34	34	35	0	0	28	30	29	29				
19	23	28	31	34	35	0	0	28	28	30	30				
24	28	29	32	34	30	0	0	28	29	30	29				
27	26	33	31	34	34	0	0	29	30	30	29				
26	33	34	31	34	34	0	0	28	30	28	30				
20	25	23	32	31	32	0	0	28	30	29	30				
12	16	26	28	26	28	0	0	29	28	29	29				
19	21	22	25	16	27	0	0	27	27	30	25	N correct	per	component	
35	35	35	35	35	35	0	0	30	30	30	30	N tot	per	component	
376	451	517	578	588	603	0	0	539	541	554	545	N correct			
665	665	665	665	665	665	0	0	570	570	570	570	N tot			
**********															

1.0		1.5		2.0		2.5		3.0		4.0		dB		
u	d	u	d	u	d	u	d	u	d	u	d	u:up	d:down	
32	26	34	32	36	35	0	0	40	39	10	10	Subject JS		
29	26	30	30	37	36	0	0	39	39	10	9			
26	24	36	34	40	40	0	0	38	40	10	10	log-complex		
32	31	37	35	38	40	0	0	39	40	9	10			
29	29	38	38	37	40	0	0	40	40	10	10			
27	26	37	38	40	40	0	0	40	40	10	10			
25	26	30	34	37	37	0	0	37	40	9	10			
31	31	38	34	37	39	0	0	39	40	10	10			
32	25	33	31	37	39	0	0	39	40	9	10			
31	32	36	40	39	40	0	0	38	40	9	10			
34	31	38	40	40	40	0	0	39	40	9	10			
32	31	38	37	39	39	0	0	40	40	10	10			
22	20	28	37	38	39	0	0	39	40	10	9			
19	20	22	27	27	36	0	0	40	39	10	10			
18	21	29	18	27	30	0	0	37	34	8	10			
25	15	21	20	23	32	0	0	29	29	5	9			
31	25	31	21	35	30	0	0	39	31	9	10			
20	27	29	29	36	30	0	0	35	36	9	10			
26	20	28	21	33	21	0	0	38	31	9	10	N correct	per component	
40	40	40	40	40	40	0	0	40	40	10	10	N tot	per component	
521	486	613	596	676	683	0	0	725	718	175	187	N correct		
760	760	760	760	760	760	0	0	760	760	190	190	N tot		
**********														
16	27	13	21	17	25	20	30	23	33	15	15	Subject NV		
20	25	21	21	16	28	25	32	25	32	15	15			
15	25	19	19	19	27	24	32	30	34	15	15	log-complex		
18	25	16	27	15	28	21	33	29	34	14	15			
14	25	24	31	24	35	29	35	35	35	14	15			
21	16	24	30	34	35	33	34	34	35	15	15			
16	28	27	29	33	33	34	33	35	35	15	15			
15	22	26	33	30	35	33	35	34	35	15	15			
17	26	26	32	34	34	35	34	35	35	15	15			
15	26	31	32	29	34	35	35	35	35	15	15			
19	28	30	34	35	33	33	35	34	35	15	15			
15	29	26	28	34	33	34	35	32	35	15	15			
19	32	29	35	33	35	35	35	32	35	14	15			
27	26	34	35	35	35	35	35	34	35	15	15			
28	21	33	33	33	35	34	35	34	35	15	15			
19	28	27	34	34	34	34	33	35	35	15	15			
12	20	24	27	30	26	35	32	33	35	14	15			
14	21	20	20	22	25	22	27	33	29	15	15			
12	21	15	20	11	16	11	17	11	20	7	8	N correct	per component	
35	35	35	35	35	35	35	35	35	35	15	15	N tot	per component	
332	471	465	541	518	586	562	617	593	637	273	278	N correct		
665	665	665	665	665	665	665	665	665	665	285	285	N tot		
**********														

1.0		1.5		2.0		2.5		3.0		4.0		dB			
u	d	u	d	u	d	u	d	u	d	u	d	u:up	d:down		
0	0	13	14	25	18	0	0	25	17	33	19	Subject AH			
0	0	9	18	18	23	0	0	24	29	37	33				
0	0	18	20	30	28	0	0	39	40	40	39	lin-complex			
0	0	22	23	38	37	0	0	39	40	40	40				
0	0	22	22	40	40	0	0	40	40	39	40				
0	0	23	25	36	40	0	0	40	40	40	40				
0	0	25	22	38	39	0	0	40	40	40	40				
0	0	15	14	35	29	0	0	38	36	40	40				
0	0	19	19	39	37	0	0	39	39	40	40				
0	0	20	13	36	24	0	0	40	34	40	39				
0	0	15	5	26	3	0	0	31	7	37	16				
0	0	8	15	20	22	0	0	28	32	36	38				
0	0	9	14	21	32	0	0	30	39	35	40				
0	0	17	15	29	30	0	0	37	38	39	40				
0	0	15	12	18	29	0	0	27	30	33	38				
0	0	15	11	26	23	0	0	30	27	31	24				
0	0	13	14	19	23	0	0	23	23	24	34				
0	0	14	11	24	23	0	0	27	31	22	38				
0	0	10	13	20	25	0	0	30	33	35	38	N correct	per	component	
0	0	25	25	40	40	0	0	40	40	40	40	N tot	per	component	
0	0	302	300	538	525	0	0	627	615	681	676	N correct			
0	0	475	475	760	760	0	0	760	760	760	760	N tot			
**********															
14	19	15	22	11	15	1	1	14	21	13	13	Subject JG			
11	18	10	14	10	17	4	4	16	21	19	16				
12	18	10	17	4	19	2	3	6	24	12	16	lin-complex			
14	25	22	33	19	25	3	5	24	30	18	20				
13	20	13	27	15	17	4	3	20	26	17	17				
22	19	25	28	23	29	3	5	24	28	19	20				
18	16	16	30	26	28	4	5	28	30	19	20				
17	20	31	26	27	27	5	4	29	30	20	20				
13	20	23	27	25	29	3	5	28	29	20	20				
13	19	23	22	17	26	4	4	24	28	19	20				
15	23	24	25	18	22	5	3	27	27	20	19				
14	23	21	21	23	25	5	5	27	30	19	20				
15	15	19	21	22	28	4	4	26	29	19	19				
17	17	29	23	25	27	5	2	30	29	20	20				
19	21	31	29	29	27	5	5	30	30	20	20				
21	17	23	19	27	23	4	5	30	29	18	20				
20	17	27	16	26	20	5	4	29	29	19	20				
22	19	26	23	29	28	5	5	30	30	20	20				
17	17	29	27	29	26	5	5	30	28	20	20	N correct	per	component	
30	30	35	35	30	30	5	5	30	30	20	20	N tot	per	component	
307	363	417	450	405	458	76	77	472	528	351	360	N correct			
570	570	665	665	570	570	95	95	570	570	380	380	N tot			
**********															

1.0		1.5		2.0		2.5		3.0		4.0		dB		
u	d	u	d	u	d	u	d	u	d	u	d	u:up	d:down	
8	16	2	24	3	22	0	0	9	26	13	33	Subject JS		
19	26	25	33	24	33	0	0	28	32	32	29			
29	26	34	33	34	35	0	0	33	35	33	34	lin-complex		
32	29	35	33	35	35	0	0	35	35	35	35			
30	30	34	32	35	35	0	0	35	35	35	35			
24	20	24	24	26	30	0	0	34	33	34	35			
19	11	22	19	24	22	0	0	28	25	32	33			
18	14	19	22	25	18	0	0	29	29	30	33			
18	20	28	23	29	27	0	0	33	34	35	34			
18	20	18	20	26	24	0	0	31	33	34	35			
22	12	24	16	28	14	0	0	29	26	32	33			
14	20	22	19	25	20	0	0	27	29	32	35			
23	16	29	19	26	20	0	0	33	30	34	34			
23	18	27	15	22	19	0	0	33	31	35	34			
17	18	24	17	22	26	0	0	24	27	33	34			
23	19	24	20	24	23	0	0	31	29	33	35			
16	14	20	14	24	23	0	0	34	31	34	34			
19	18	22	20	28	27	0	0	33	31	34	34			
24	20	28	21	27	25	0	0	33	28	35	35	N correct	per	component
35	35	35	35	35	35	0	0	35	35	35	35	N tot	per	component
396	367	461	424	487	478	0	0	572	579	615	644	N correct		
665	665	665	665	665	665	0	0	665	665	665	665	N tot		

****

0	0	21	15	15	20	2	4	26	27	33	33	Subject NV		
0	0	15	28	29	27	5	5	33	31	34	35			
0	0	16	25	31	34	4	5	34	35	34	35	lin-complex		
0	0	24	28	31	33	5	5	35	35	35	35			
0	0	26	29	31	33	5	5	34	35	34	35			
0	0	21	17	24	28	5	5	34	33	35	35			
0	0	22	25	29	35	5	5	34	34	35	35			
0	0	30	33	29	33	5	5	35	35	35	35			
0	0	29	31	30	34	5	5	35	35	35	34			
0	0	23	20	26	30	5	5	35	33	35	35			
0	0	19	21	25	25	4	5	33	34	35	35			
0	0	22	20	23	22	3	5	32	33	35	33			
0	0	18	16	18	27	3	4	32	35	34	35			
0	0	20	18	20	25	2	5	32	24	35	30			
0	0	17	17	14	17	3	0	20	30	26	35			
0	0	16	15	13	17	3	2	19	24	28	31			
0	0	18	16	17	18	2	4	8	26	28	34			
0	0	18	16	14	26	2	4	23	29	29	34			
0	0	20	9	9	21	1	3	15	22	12	21	N correct	per	component
0	0	35	35	35	35	5	5	35	35	35	35	N tot	per	freq
0	0	395	399	428	505	69	81	549	590	607	635	N correct		
0	0	665	665	665	665	95	95	665	665	665	665	N tot		

****