# Discriminality of statistically independent Gaussian noise tokens and random tone-burst complexes

*Citation for published version (APA):*
Goossens, T. L. J., Par, van de, S. L. J. D. E., & Kohlrausch, A. G. (2007). Discriminality of statistically independent Gaussian noise tokens and random tone-burst complexes. In B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, & J. Verhey (Eds.), *Hearing - From sensory processing to perception* (pp. 343-351). Springer. https://doi.org/10.1007/978-3-540-73009-5_37

*DOI:*
10.1007/978-3-540-73009-5_37

*Document status and date:*
Published: 01/01/2007

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

***Please check the document version of this publication:***

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

Link to publication

*General rights*
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.tue.nl/taverne

*Take down policy*
If you believe that this document breaches copyright please contact us at:

openaccess@tue.nl

providing details and we will investigate your claim.

Download date: 04. Oct. 2023

# 37 Discriminability of Statistically Independent Gaussian Noise Tokens and Random Tone-Burst Complexes

TOM GOOSSENS[1], STEVEN VAN DE PAR[2], AND ARMIN KOHLRAUSCH[1,2]

## 1 Introduction

Hanna (1984) has shown that noise tokens with a duration of 400 ms are harder to discriminate than noise tokens of 100 ms. This is remarkable because a 400-ms stimulus potentially contains four times as much information for judging dissimilarity than the 100-ms stimulus. Apparently, the ability to use all information in a stimulus is impaired by some kind of limitation, e.g. a memory limitation (cf. Cowan 2000) or a limitation in the ability to allocate attentional resources (cf. Kidd and Watson 1992).

In a first experiment, this study examined the influence of stimulus duration and bandwidth of Gaussian noise tokens on the ability to perform an auditory discrimination task. In a second experiment, the amount of potential information in a stimulus was decoupled from its duration in order to more carefully examine the properties of the memory or attention limitation that results in the discrimination impairment. Finally, a computational model that limits the amount of perceptual information is introduced as an attempt to model the findings of the first and second experiment.

## 2 Discrimination of Gaussian Noise Tokens

### 2.1 Method and Stimuli

This psychoacoustic experiment is a replication of, and partly an extension to, an experiment by Hanna (1984). It was executed to test the ability of listeners to discriminate between Gaussian noise tokens. The experiment was performed using a *same-different* procedure where two noise tokens were presented to the listener in each experimental trial. These noise

---

[1] University of Technology Eindhoven, The Netherlands, t.l.j.goossens@tm.tue.nl

[2] Philips Research Laboratories Eindhoven, The Netherlands, steven.van.de.par@philips.com, armin.kohlrausch@philips.com

Hearing – From Sensory Processing to Perception
B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey (Eds.)

tokens were either identical or uncorrelated. For each trial, new noise samples were generated. The a priori probability of *same* and *different* presentations was 50%. Subjects were given feedback about the correctness of their answer after each trial.

Three subjects participated, including the first and second author, in the experiments that were divided in several sessions of maximally 1 h. Each experimental condition was measured in four blocks of 50 (Subjects S1 and S2) or three blocks of 100 (subject S3) trials. The blocks were presented in random order. Each subject had a training session of at least 25 blocks of 100 trials before the actual experiment started.

The responses of the listeners were transformed into $d'$ values by calculating percentages correct of same and different presentations for each condition. These percentages correct were converted to $z$-scores. Finally $d'$, the quantity of interest, was calculated by adding the $z$-scores of same and different presentations. At chance performance, the $d'$ value equals zero. Above-chance performance results in positive $d'$ values, e.g. 69% and 84% correct for both same and different trials results in $d'$ values of approximately 1 and 2, respectively.

Measurement conditions were combinations of five noise bandwidths and nine durations. The −3 dB bandwidths were: 100–3300 Hz; 100–600 Hz; 225–275 Hz; 2800–3300 Hz; and 2975–3025 Hz. The specified durations before filtering were: 1.6, 6.4, 10.2, 16.1, 25.6, 40.6, 64.5, 102.4 and 409.6 ms.

Noise tokens were produced by digitally generating broadband noise of the specified duration with 40 dB/Hz SL. Subsequently, the tokens were filtered with a Chebyshev Type II digital filter with slopes of at least 100 dB/oct. The stimuli included the ringing of the filters to avoid audible truncation effects.

The stimuli were presented from a PC through a high-quality soundcard at 16 bit, 44.1-kHz sampling resolution on Beyerdynamic DT990Pro headphones.

### 2.2 Results and Discussion

Figure 1 shows the results of the Gaussian noise token discrimination experiment. The abscissa indicates the noise duration and the ordinate indicates the across-subject means of the $d'$ values. In order to get an estimate of the confidence intervals, a bootstrap procedure was used to create subsets of repeated measurement blocks across all subjects. The means of these subsets were used to calculate means and 95% confidence intervals. Bootstrap sample size was 1000.

Looking at the influence of bandwidth on the results, it can be observed that for the stimuli comprising low frequencies (dashed lines and solid line), discrimination ability improved with increasing bandwidth. For high frequencies (dash-dotted lines) discrimination proved to be very difficult for subjects in all conditions.

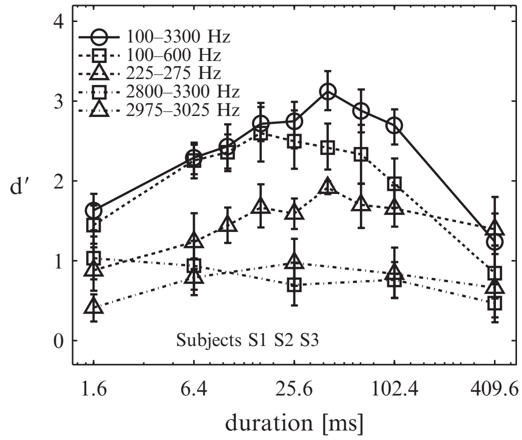

**Fig. 1** Across-subject *d′* means for Gaussian noise token discrimination of three subjects for broadband conditions (*circles*), 500-Hz wide bandpass conditions (*squares*), and 50-Hz wide bandpass conditions (*triangles*) at low frequencies (*dashed lines*) and at high frequencies (*dash-dotted lines*). The *abscissa* indicates the noise duration. *Error bars* indicate 95% confidence intervals

With increasing duration, discrimination ability increased up to a maximum performance at durations of about 40 ms. Beyond this maximum, the ability to discriminate decreased with increasing duration. This degradation of discriminability is remarkable because the longer duration stimuli contain more information for performing discrimination. These results were similar to the results found by Hanna (1984). It seems that listeners do not have access to all information that is available in the stimulus. In the next experiment the duration effect will be studied more carefully.

## 3 Discrimination of Random Tone-Burst Complexes

In the previous experiment, the number of degrees of freedom in the stimulus increased with increasing duration. In this experiment the number of degrees of freedom is decoupled from the stimulus duration.

### 3.1 Method and Stimuli

The method used for this second experiment was the same as for the previous experiment – only the stimulus type was different. Instead of Gaussian noise tokens, random tone-burst complexes were used as stimuli.

The random tone-burst complexes in this experiment consisted of a number of 5-ms Hanning-windowed tone-bursts placed at random time positions

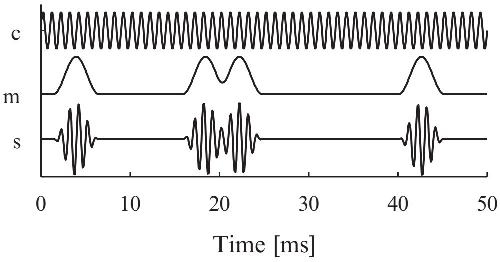

**Fig. 2** A random tone-burst signal (s) is generated by multiplying a carrier (c) with a modulation envelope (m) that consists of a number of Hanning windows additively placed at random temporal positions within the duration of the stimulus

within a time-frame of either 51.2 ms or 409.6 ms (cf. Fig. 2). This was done by multiplying a 70-dB SPL carrier tone (trace *c* in Fig. 2) of appropriate duration and frequency with a modulation envelope (trace *m* in Fig. 2). This modulation envelope was generated by adding a number, depending on the condition, of Hanning-windows at random temporal positions within the stimulus duration. This way of stimulus generation did not introduce phase mismatches when two tone bursts of the same frequency overlapped.

A random tone-burst complex allows for distribution of the amount of spectral and temporal information in the stimulus in a duration-independent way whereas the amount of information in a Gaussian-noise token always increases with increase of duration.

In one set of conditions, the tone bursts all had a frequency of 607 Hz (ERB rate of 12). In another set of conditions there were tone bursts of seven frequencies, i.e. 208, 314, 444, 607, 808, 1057, and 1367 Hz (ERB rates of 6 up to and including 18 with a spacing of 2). In both sets, the number of tone bursts was 2, 4, 8, 16, 32, 64, 128, and 256 tone bursts *per frequency*, with the exception that for the 51.2-ms conditions, the 128, and 256 tone bursts per frequency were not measured.

Per experimental condition, trials were presented in 4 blocks of 100 trials. Blocks were presented in random order in sessions of maximally 1 h to four subjects, including the three subjects of the first experiment. Each subject had a training session of at least 8 conditions of 100 trials before the actual experiment started.

### 3.2 Results and Discussion

The results of the random tone-burst discrimination experiment are shown in Fig. 3. The across-subject means of the $d'$ values are indicated on the ordinate. The error bars show the 95% confidence intervals. Again, a bootstrap procedure was used to calculate the means and 95% confidence intervals.

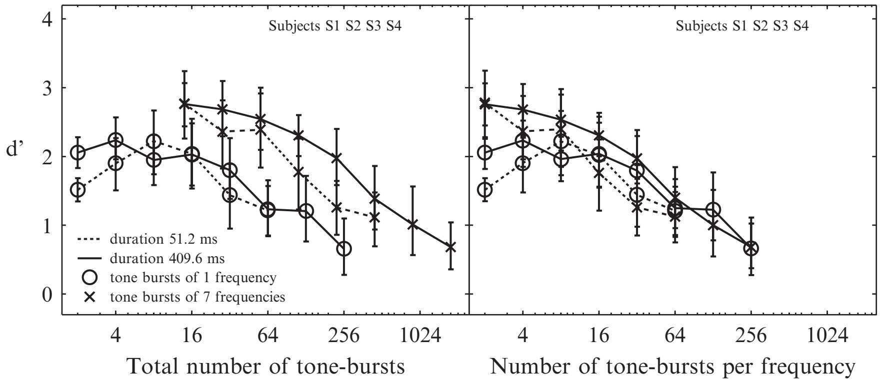

**Fig. 3** Across-subject $d'$ means for random tone-burst discrimination of four subjects. *Error bars* indicate 95% confidence intervals

The left and right panel show the same results only represented on different abscissas. The left panel shows the results as a function of the *total* number of tone bursts, whereas the right panel shows them as a function of the number of tone bursts *per frequency*. Note that the conditions are much more alike in the right panel implying that the total number of tone bursts is less predictive for the ability to discriminate than the number of tone bursts per frequency. However, it has to be noted that this trend is less clear for individual subjects than in the average data.

All conditions show the same trend except that the single frequency conditions (circles) with two tone bursts show worse performance than the conditions with four tone bursts. Again, this means that more potential information in a stimulus does not necessarily improve discrimination performance.

Moreover, in contrast to the first experiment, the duration of the stimulus did not show a strong effect on performance. This supports the assumption that the cause of the degradation of performance for longer durations in the first experiment lies in a limitation of the amount of information that can be kept in memory rather than a limitation related only to stimulus duration.

## 4 A Discrimination Model

### 4.1 Method

Several psychoacoustic models have been proposed for predicting the discrimination of stimuli (e.g., Dau et al. 1996). These make use of an internal representation (IR) of a stimulus. It is inherent to the IR approach that the number of degrees of freedom of the IR increases with the duration of the stimulus; i.e. there is more information. According to such models, long duration stimuli are easier to discriminate than short duration stimuli, which contradicts the findings in the first experiment of this paper and those of Hanna (1984).

In this chapter a model is introduced that fixed the number of degrees of freedom in the IR. It consisted of three stages of which the *first stage* computed an IR of both stimuli in a trial according to the model by Dau et al. (1996). This model includes outer and middle-ear filtering (not used in the present simulations), basilar-membrane filtering (gamma-tone filter-bank, 52 filters from 20 Hz to 10 kHz), inner hair cell, adaptation, temporal smoothing, and addition of internal noise. The number of degrees of freedom of this IR was still dependent on the duration.

The *second stage* reduced the degrees of freedom to a fixed amount regardless of the duration of the stimuli. This was done by simply windowing each auditory channel of the IR with 75%-overlapping Hanning windows of a length of 1/3 of the stimulus duration, including 150 ms ringing of the filters, and by taking the average IR across each window. Note that by doing so, the window length of the Hanning windows is dependent on the stimulus duration and thus the number of values obtained in each auditory channel is fixed. The window averages together form a new set of values, or points, which essentially is a reduced fixed-size IR.

The *last stage* was a decision device that decided if two presented stimuli in a trial were the same or different. For this purpose it first calculated a decision variable by taking the sum of squares of the difference in IR integrated across time and frequency. Then, a decision noise was added. Finally, the decision variable was compared to a certain criterion. If it was larger than the criterion, then the decision device decided that the stimuli were different, else it decided they were the same.

The criterion was determined heuristically. At the start of a block of 100 trials, the criterion was set to a fixed arbitrary value. However, this criterion was adjusted after each trial by storing the values of the decision variable in two separate storages. One storage for values that, after feedback from the experiment, were the same and the other for values that were different. By making Gaussian fits, the decision device calculated means and variances for both cases to determine a new criterion using maximum-likelihood estimates. In every subsequent trial, the model adapted to a more accurate criterion and its performance improved. After about six to ten trials the criterion had converged to a reasonable value.

The variance of the memory noise was chosen such that the sum of squares of the difference between the model simulations and the experimental data, expressed in $d'$, was minimal. Different variances of the observation noise were needed for the simulation of the first and second experiment. This reflects the different nature of the stimulus features in both experiments.

Per condition, trials were presented to the model in 8 blocks of 100 trials.

## 4.2 Results and Discussion

The left panel of Fig. 4 shows the results of the model simulations of the first experiment. As can be observed when comparing this panel to Fig. 1, the order of the bandwidth dependencies was correctly predicted by the model, although it showed too strong spectral integration, especially for shorter durations. For example, in Fig. 4 (left panel) the simulated broadband conditions (circle markers) do not coincide with the simulated low-frequency 500-Hz conditions (dashed line, square markers) in the way that they do in the psychoacoustic data.

The effect of too much spectral integration can also be observed in the right panel of Fig. 4 where the results of the model simulations of the second experiment are shown. It is demonstrated by the fact that the seven-frequencies conditions of the 51.2-ms conditions (dotted line with cross markers) did not coincide with the corresponding 409.6-ms conditions (solid line with cross markers) as they did in the psychoacoustic data shown in Fig. 3. In contrast, these two curves do coincide in the single-frequency conditions where no frequency integration was possible.

The model correctly predicts the dependence of discriminability on stimulus duration for the first experiment showing a maximum at about 40 ms. This model behavior is related to the fixed number of points in the fixed-size IR that is independent of stimulus duration. As a result, the discriminability of the stimulus now only depends on the variance of each of the points in the fixed-size IR. When the stimulus duration is about 40 ms, it has a number of degrees of freedom comparable to the number of points in the IR and as a result the variance in the internal representation is large, resulting in a high discrimination performance of the model. For shorter-duration stimuli, fewer degrees of freedom will be present in the stimulus and the points within the internal representation will be highly correlated,

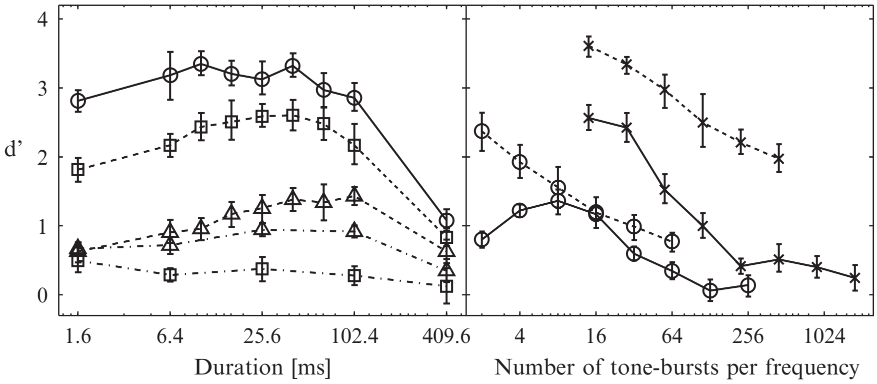

**Fig. 4** Model results for the data of Fig. 1 (*left panel*) and the data of Fig. 3 (*right panel*)

reducing the discriminability. For longer-duration stimuli, the number of degrees of freedom will be much larger than the number of points in the IR and averaging will take place reducing the variance in the points of the fixed-size IR and reducing discriminability.

A final remark about the model that should be addressed is the fact that, as analysis has shown, the variability in the fixed-size IR mostly takes place at the on- and offset of the stimulus. For this implementation of the model it would be preferable to have more homogeneous variability, as we know from other studies, e.g. Fallon and Robinson (1992), that listeners are able to discriminate based on information in the middle of a stimulus as well.

## 5 General Discussion

The simulations show that a simple method for limiting the amount of information in the internal representation can already account for a large portion of the effect of degraded discriminability in the experiment with noise tokens of various durations. It should be noted however, that the *method* of limiting the amount of available information presented in this study should not be taken as a definitive method of memory limitation. The exact nature of this limitation needs further investigation.

There is a difference in the processing of spectral and of temporal information. While increasing the amount of temporal information of a stimulus can lead to impairment of the ability to discriminate, increasing the amount of spectral information has never led to such impairments in this study.

The random tone-burst experiment, in which the amount of information was decoupled from duration, has shown that the ability to discriminate is not so much a function of the duration but rather of the number of degrees of freedom or the amount of information in a stimulus.

## References

Cowan N (2000) The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav Brain Sci 24:87–185

Dau T, Püschel D, Kohlrausch A (1996) A quantitative model of the “effective” signal processing in the auditory system. I. Model structure. J Acoust Soc Am 99:3615–3622

Fallon SM, Robinson DE (1992) Discrimination of burst of reproducible noise. J Acoust Soc Am 92:2630–2635

Hanna TE (1984) Discrimination of reproducible noise as a function of bandwidth and duration. Percept Psychophys 36:409–416

Kidd GR, Watson CS (1992) The “proportion-of-the-total-duration rule” for the discrimination of auditory patterns. J Acoust Soc Am 92:3109–3118

## Comment by Emiroğlu

Did you try to “squeeze” the temporal dimension in the model, i.e. to average the “Internal Representation” over all time-steps or over a certain minimal duration, respectively, before feeding the IR into the decision device (i.e. optimal detector)? Because I think, the brightness (or “brightness melody” if the stimulus is long enough) is here the distinguishable and memorizable cue, and the “spectral percept” probably needs a minimal time span that is necessary to establish/recognize one “brightness”.

## Reply

Two alternative approaches to reduction of the information in the internal representation (IR) are addressed in your question.

The first approach is to collapse the time dimension of the IR by taking the average over its full duration, resulting in a one-dimensional IR with only spectral information in it. Simulations show that the duration at which the ability to discriminate is maximal, is inversely related to the number of (temporal) degrees of freedom in the IR. Hence, averaging across the complete duration does not improve the predictions of the model.

The second approach is to take a fixed window-length (called minimal duration in your comment) instead of making the window length dependent on the stimulus duration. This would cause the degrees of freedom in the IR to increase with increasing duration. As an effect, the ability to discriminate would also increase and this is not what we observe in the measurements.