The relation between rhythm perception and production: towards a Bayesian model

Makiko Sadakata[1], Peter Desain[1] and Henkjan Honing[1,2]

Music, Mind, Machine Group, NICI, University of Nijmegen[1]
P.O. Box 9104, 6500 HE Nijmegen, The Netherlands

ILLC / Music Department, University of Amsterdam[2]
Spuistraat 134,1012 VB Amsterdam, The Netherlands

E-mail: m.sadakata@nici.kun.nl, Desain@nici.kun.nl, honing@hum.uva.nl

Abstract The aim of this study is to shed light on the relation between rhythm perception and production. We try to show that empirical results concerned with rhythm perception and production described in the literature are in fact related according to Bayes-rule. This is addressed in a meta-analysis comparing empirical results from a number of existing perception and production studies.

Keyword rhythm perception, rhythm production, Bayesian model

1. Introduction

The aim of this study is to shed light on a new explanation of the relation between rhythm perception and production. Since processing of time intervals plays an important role in our daily life, the importance of studying time relations as a mental phenomenon is already clearly brought up at the end of the 19th century (Jastrow, 1890).

We will test three hypotheses concerning the relation between rhythm perception and production. The first states that perception and production make use of the same mental representations and behave the same. Much classical work has been based on this assumption (e.g., Eisler 1976). However, few recent studies claim the relationship to be this simple. They support a second hypothesis that perception and production stem from different processes. There is some evidence in the psychophysical literature for this (e.g., Sternberg, 1982) showing different identification profiles in perceptual versus production tasks. However, Bayesian inference might be able to give a new way to interpret the empirical data, supporting a third hypothesis that perception and production seem different, but this difference is actually the result of how the rhythmic categories are sensitive to competition in perception.

2. Bayesian approach

2-1. Motivation for Bayes' theorem

The hypothesis that biological perceptual systems can be explained using a Bayesian approach has been tested in the field of visual perception to a great success. The underlying view of this research is that "the Bayesian formulation captures the essence common to most of the frameworks, and allows the distinctions to be articulated clearly (Knill & Rechards, 1996)".

In that sense we might be able to adopt Bayesian inference into the nature of human processing for temporal patterns as well.

2-2. Conceptual frameworks

Rhythm production data is interpreted as a probability of a performance given a score category. Likewise, perceptual data is interpreted as the probability of a category given a performance. The relation between these aspects is as follows: the production data is characterized as a probability distribution that is the relative likelihood of the perceptual process. And so, perceptual data can be predicted as a posterior distribution, which is determined in part by production data, including the nature of the noise added in the production, and in part by the statistical properties of scores. Bayes' rule specifies a way to partially decompose the posterior into these parts. According to Bayes' rule, the posterior is given by

$$p(category \mid performance) = \frac{p(performance \mid category) \times p(category)}{p(performance)}. \quad (1)$$

Fig. 1 shows example distributions of the perception (A) and production (B) of three different rhythmic categories, vertical axes show probability and horizontal axes show duration of the first interval of the categories respectively. The posterior distributions can be reinterpreted into their independent distributions and their priors using Bayes-rule, extracting the effect of competition.

We will accomplish the process through the following steps.

(i) Data approximation
(ii) Predicting perceptual data from production data using Bayes' theorem
(iii) Comparison

3. Data descriptions

To be able to yield a relevant comparison across situations with a different method, we did a careful selection of the data sets. All of the data sets needed to use rhythmic patterns consisting of two intervals whose total duration was 1000 ms. Some of the data sets required interpolation or extrapolation. Detailed procedures of experimental procedures can be found in the publications.

Although our whole project will be addressed in a meta-analysis comparing empirical results from a number of existing perception and production studies, here we demonstrate the method using only two data sets.

3-1. Perceptual data

In Desain & Honing (in press) a categorization experiment is described for three interval patterns. The same set up was used

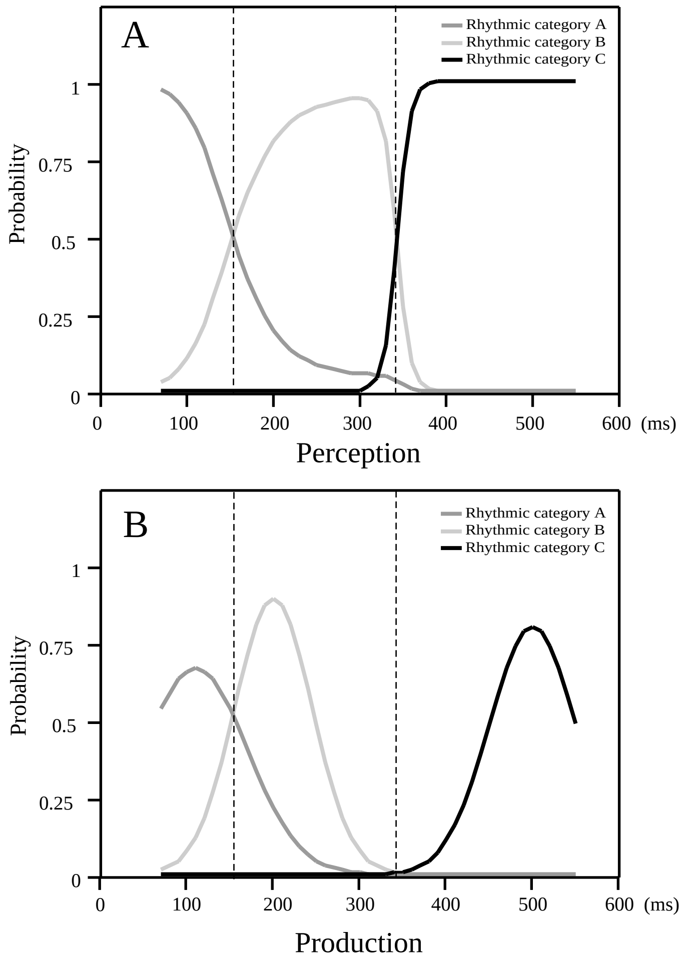

Fig. 1 A: Example distribution of perception with competition of the categorical boundaries, given as formula (1).
[$p(category \mid performance)$]
B: Example distribution of production with no competition of the categorical boundaries multiplied by priors, given as the numerator of the formula (1).
[$p(performance \mid category) \times p(category)$]

to collect two interval categorizations. Each stimulus pattern was made up of two time intervals on a time grid of 1/19th of 1000 ms, the minimum duration of an interval being three time grid units and the maximum 16 units. The task was to notate the stimuli in common musical notation.

3-2. Production data

Sadakata et al. (2002) deals with production data performed by twelve percussionists. The task was to perform the rhythmic patterns consisting of two intervals whose duration ratios were 1:1, 1:2, 1:3, 1:4, 1:5, 2:1, 3:1, 4:1 and 5:1. We selected the tempo condition 60 (total duration of the two intervals is 1000 ms). However, though close, not all observations of production data in this condition sum to exactly 1000 ms due to the nature of this kind of task. Therefore observations were normalized.

4. Analysis

4-1. Data approximation

The Beta distribution was chosen for the data approximation because of its' flexibility, as it can fit skewed distributions. It describes a family of curves that are nonzero only on the interval [0, 1]. It has two free parameters, α and β, which characterize the form of curve, plus two parameters to rescale and shift the distribution to any mean *m* and width *w* by using a linear transformation. The parameters α, β, *w*, and *m* were estimated from the production data set for every rhythmic pattern, using the maximum likelihood method in JMP (ver. 5.0). The approximated data and the histogram of the raw data are shown in Fig. 2.

We also approximated the data using a normal distribution, whose parameters can be simply calculated as mean and standard deviation of the data. An example of these two approximations is shown in Fig. 3. The vertical axis shows probability and horizontal axis shows duration of first interval, respectively. The KSL (Kolmogorov-Smirnov Lillifors) goodness of fit test revealed that generally the beta distribution provides a better approximation than the normal distribution.

4-2. Data processing

Perceptual data is shown in Fig. 4 as distribution *a*. Vertical axis shows probability and horizontal axis shows duration of the first interval on a time grid of 1/19th of 1000 ms. Rhythmic categories in this data set, not occurring in the production data, such as 2:3, were removed. Because their contribution was quite small, they could be removed, and the data re-normalized, without a large change.

Production data (without any transformation) is shown in Fig. 4 as distribution *b*, presented on the same grid of distribution *a*.

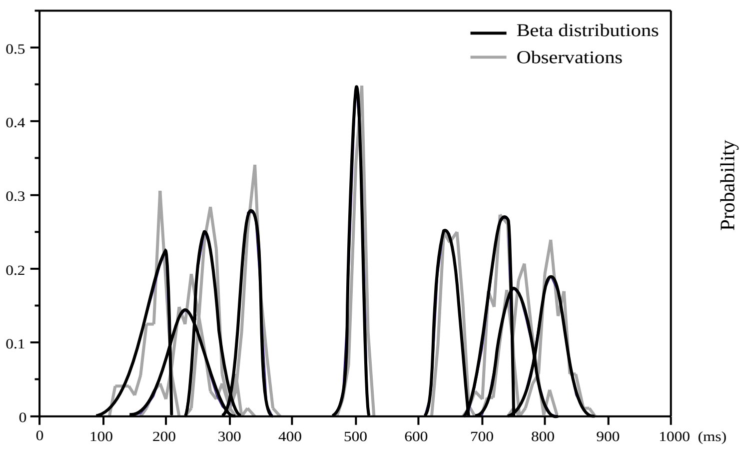

Fig. 2 The histogram of the observations (gray line) and the approximating beta distributions (black line).

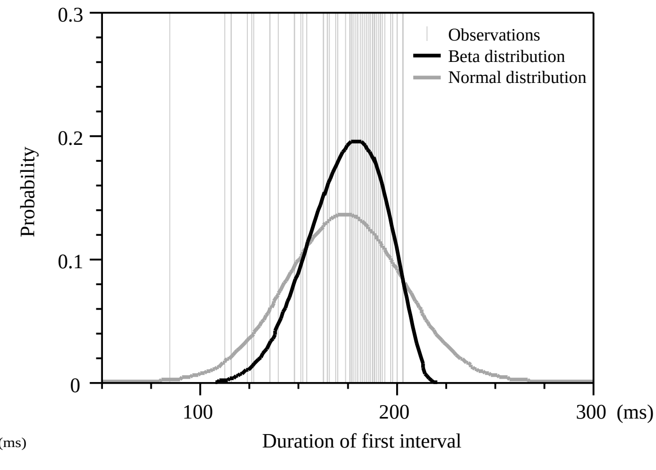

Fig. 3 Relation between observed data and two approximated continuous distributions. Gray vertical lines show each of the observations, gray curve represents data approximated by the normal distribution and black curve line represents the data approximated by the beta distribution.

Probabilities were calculated on the time grid as the perceptual data. By substituted known quantities (production data) for likelihood, we will get the predicted posterior probabilities. The probability of a certain perceived category at the certain time grid will be predicted as (total number of the rhythmic categories are nine)

$$Perception_i = \frac{\Pr ocudtion_i \times \Pr ior_i}{\sum_{j=1}^{9} \Pr oduction_j \times \Pr ior_j}. \qquad (2)$$

The predicted perceptual data using uniform priors, assuming all categories to be as likely, are shown in Figure 4*c*.

As we already know the data for both of rhythm perception (posterior) and production (likelihood), we can estimate the priors, fitting the predicted perception data to the observed, minimizing the root mean square (rms) error between them. The predictions with the optimal estimated priors are show in Fig. 4*d*.

Note that the formula (2) requires overlap between the rhythmic categories in production: the sampling of categories needs to be dense enough. This is not the case around the 1:1 ratio. We calculated the distributions separately for the left (1:5, 1:4, 1:3, 1:2) and the right part (2:1, 3:1, 4:1, 5:1). Between them the boundaries between categories are predicted by formula (2). The left and right boundary of the 1:1 category cannot be checked by the present choice of datasets.

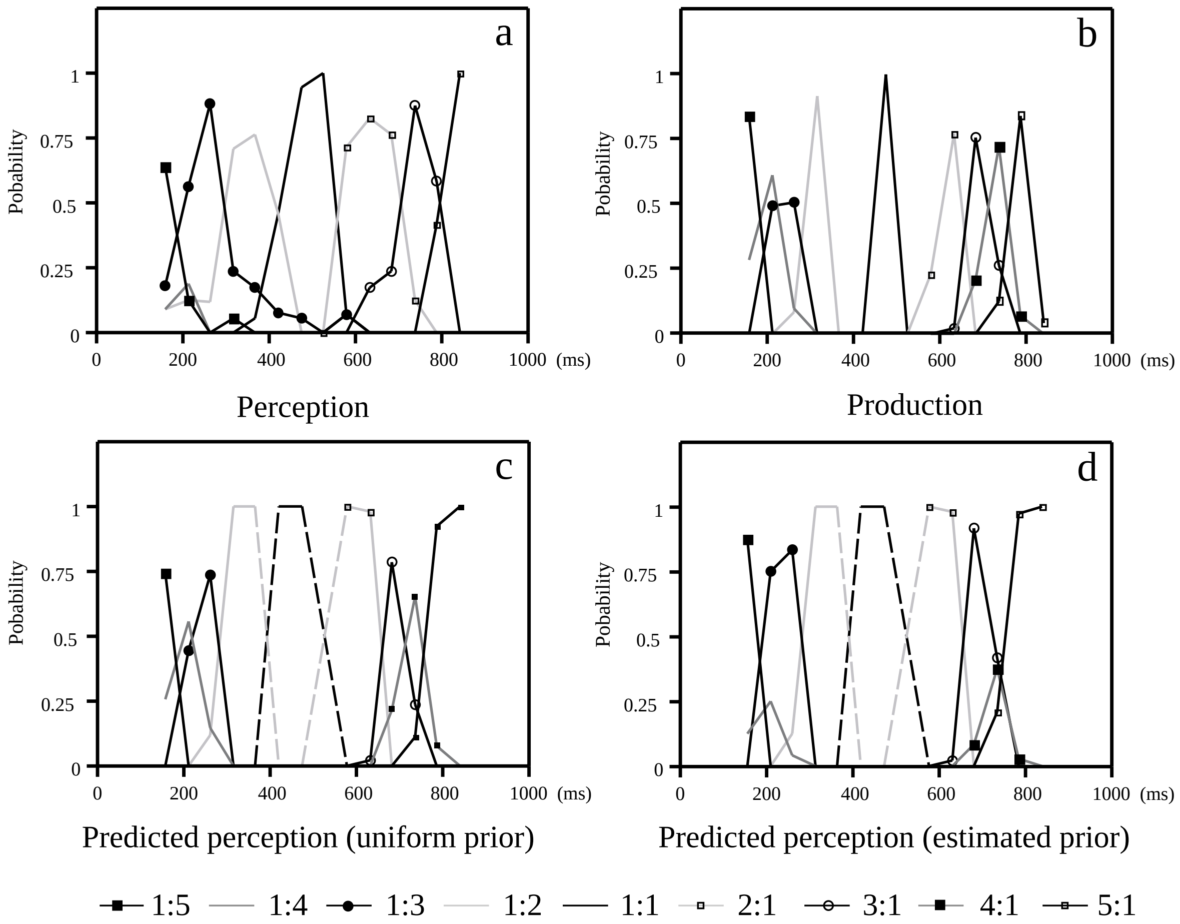

Fig. 4 The observed and predicted distributions of nine rhythmic categories. On the horizontal axis the first time interval is given, on the vertical axis the probability is represented, either of producing this interval given a rhythmic category, or of judging this interval as a proper representation of the given rhythmic category.

The distance from distribution *a* to each of three distributions (*b*, *c* and *d*) was calculated as the rms error. For *b* this was 1.88, for *c* 1.70 and *d* 1.33. This indicates that distribution *d* is the one most close to distribution *a*.

5. Discussion

Hypothesis 1 and 2 can be examined by inspecting the relation between distribution *a* and *b*, which is a direct comparison of rhythmic perception and production. It is clear that distribution *a* and *b* are far from being identical, which means that hypothesis 2 may be more close to the truth than hypothesis 1. However, it is also clear that distribution *c* and *d* are closer to distribution *a* than *b* is. This is evidence for hypothesis 3, and supports the claim that the consistency of human processing for temporal pattern is indeed constrained by Bayesian rules.

One may wonder the role of the estimated priors that were used in obtaining distribution *d*. We may be able to interpret these as a quantity that reflect the characteristics of the rhythmic categories, such as simplicity or familiarity. In order to pinpoint more precise the role of the priors and broaden the support for the validity of the Bayesian approach, further research using other data sets (Repp et al., 2002; Sternberg et al., 1982) is currently in progress.

References

[1] P. Desain & H. Honing, The formation of rhythmic categories and metric priming, Perception, In press.

[2] H. Eisler, Experiments on Subjective Duration 1868-1975: A Collection of Power Function Exponents, Psychological Bulletin, vol. 83, No. 6, pp. 1154-1171, 1976.

[3] J. Jastrow, The Time-Relations of Mental Phenomena (Book Review), Mind: a quality review of psychology and philosophy, vol. 16, pp. 545, 1890.

[4] D. C. Knill & W. Richards, ed., Perception as Bayesian Inference, pp. x, Cambridge University Press, Cambridge, 1996.

[5] B. H. Repp, W. L. Windsor and P. Desain, Effects of tempo and timing of simple musical rhythms. Music Perception, 19(4), 563-591, 2002.

[6] M. Sadakata, K. Ohgushi & P. Desain, A cross cultural comparison study of the production of simple rhythm patterns, Proc. Of ICAD 2002 Rencon work shop, pp. 24-27, Kyoto, Japan.

[7] S. Sternberg, R. L. Knoll & P. Zukofsky, Timing by Skilled Musicians, In D. Deutsch (Ed.), The psychology of music, pp. Academic press, London, pp.181-239, 1982.