STM 1978:1

RHYTHMSYVARD

A computer program for analysis of rhythmic performance

By Ingmar Bengtsson, Alf Gabrielsson and Barbro Gabrielsson

RHYTHMSYVARD

A computer program for analysis of rhythmic performance

By Ingmar Bengtsson, Alf Gabrielsson and Barbro Gabrielsson

1. Introduction

The rhythm research project in Uppsala has been presented before in this journal (Bengtsson, Gabrielsson&Thorsén, 1969) as well as in other papers, notably Bengtsson (1974) and Bengtsson & Gabrielsson (1977). Its two main aims are to investigate *rhythm as performed* and *rhythm as perceived.*

Concerning rhythm as performed—within the project so far studied only in music with metrically regular patterns (Western art music, popular, and folk music)—a central hypothesis is that "an important stimulus factor for the eliciting of various rhythm characters is to be found in some kind of *systematic variations as regards tone durations* (SYVAR-D)" (Bengtsson & Gabrielsson, 1977, p. 30). In particular, an important part of the project deals with "patterns" that are typical of "special ways of playing", "the right way of playing" etc., e.g. according to some stylistically, regionally or otherwise defined rhythmic "dialect". At least in such performances, SYVAR-D patterns have been shown to exist in all cases investigated to date.

In order to facilitate such investigations, a computer program, RHYTHM-SYVARD, has been developed. A first, basic version was designed by Bernt Castman and Lars Elve Larsson at the Uppsala University computer centre (Castman, Larsson & Bengtsson, 1974). A second, more complete version was written by Sverker Johansson in 1976. A brief description of the latter and its use is given below. A detailed technical manual is as yet available only in Swedish.

2. Raw data (D values)

Any collection of duration values (D values) from a musical performance, no matter how they are obtained, may serve as raw data for processing by the program. In the following, we use D values obtained from graphic registrations made by the MONA analyzer (Bengtsson, Gabrielsson & Thorsén, 1969; Bengtsson, 1967a; Bengtsson, Tove & Thorsén, 1972).

In series of musical sound events (SE), one must usually distinguish between three types of "durational content": (a) the duration from the onset of a sound to its end, here called *duration in—out,* abbreviated *Dio,* (b) the duration from the onset to the onset of the next sound, i.e. *duration in—in* or *Dii,* and (c) the duration from the end of a sound to the onset of the next sound, *duration out—in* or *Doi.* In music, the Dii values are usually considered to be particularly important. Doi values may represent real rests (and may appear as such in the notation), or be the result of articulation, or just of rapidly decaying sounds as on plucked string instruments and many percussion instruments.

In the RHYTHMSYVARD program, the investigator is free to keep a Dio value (say, 200 milliseconds) and an immediately following Doi value (say, 70 ms) separate, or to treat them as a Dii value (270 ms). Such choices are made by means of the Numericode labelling (see section 4 below).

A simple Swedish tune will be used as an example throughout this text, see *Fig. 1.* The tune consists of 30 tones. If it is repeated directly, or a repeat is hinted at by the playing of at least the first tone, the final tone (no. 30) can be measured. If not, it is not possible to obtain an exact D value for this SE (at least not a Dii value), and it must therefore be disregarded.

3. On terminology

It was found necessary to develop a special conceptual framework and terminology, both for classification and labelling purposes when using the program and to describe metrical and other features of the performed musical structures. It is desirable, in the latter case, to find designations that are not tied too closely to notational concepts and terms (e.g. "dotted" rhythms or notes, since "dots" are never perceived). To facilitate the understanding and the use of the RHYTHMSYVARD program, the most important of these concepts, labels, and designations will be briefly mentioned.

(*a*) One *performer* or several performers have produced, each of them or together, one or several renderings of a certain piece of music. "Piece of music" should be understood in a very wide sense; it could, for instance, consist of only a series of beats on a drum. The music must be easily identifiable as regards pitch order, duration relationships etc. (in the way that such parameters are usually shown in ordinary notation). However, the "piece of music", here called MELODY, may comprise several *variants,* differing in details with regard to pitches, number of tones, and time values. From these facts a simple hierarchy can be derived and used as a classificatory system with four levels:

1) MELODY = a piece of music as defined above,

2) VARIANT = a particular variant (if any) of the same MELODY,

3) PERFORMER = a certain person or group of persons who has performed a MELODY/VARIANT, and

4) VERSION = the particular rendering of a certain MELODY/VARIANT by a certain performer or certain performers.

If a certain rendering includes a full repeat of a MELODY/VARIANT, there is a free choice between treating the repeat as a part of the VERSION or as another VERSION.

The levels are labelled in the order just stated with four numerals separated

by dots. This is the *identificator* of a particular version. Thus 7.1.3.2 signifies MELODY no. 7 (within a certain larger material of several melodies), VARIANT no. 1, PERFORMER no. 3, and his/her VERSION no. 2.

All versions that are treated together in a particular running of the RHYTHMSYVARD program constitute what is called a "total object", TOTOBJ.

(*b*) Another set of levels concerns *metrical formats,* from a single sound event to a musical period or a whole MELODY. Here the following levels and designations have been chosen:

E (element) or SE (sound event, including possible Doi values) = each single tone or rest that is represented by its own measured D value (and its Numericode label, see below);

B = beat, i.e. the class of D values that clearly represents, or is chosen to represent, the behavioural and/or perceptual beat level in the MELODY/VERSION;

M = measure; the M format may or may not agree with some notated version of the piece, depending on the investigator's choice;

M groups = groups of measures, generally starting with odd M numbers, and labelled 2M-GR, 4M-GR etc.;

PERIOD = a closed form structure, in notation ordinarily or ideally of the size of about 8 or 16 measures; in the RHYTHMSYVARD program, PERIOD means any structure of about that size selected by the investigator.

Order numbers are shown by a numeral *after* the letter: M3 = measure no. 3, B4 = beat no. 4, and so on. *Number of metric entities* on a certain level is shown by a numeral *before* the letter: 4B = four beats, 2M = two measures etc. M/2 designates half a measure, M/3 a third of a measure, etc.

Position within the overall metrical structure, and thus in several measures, may be signalled by the letter p before the main letter. Thus pB2 means all beats in the position "B2" within a certain MELODY/VERSION.

Time signatures are written as in ordinary notation (i.e. as 3/4 in our example); however, C must be written 4/4, and ₵ as 2/2 (see the Numericode labelling system under 4 below).

Pairs of D values may be of three kinds: the two SE may be equal (designated EQ), the first may be longer and the second shorter (LS), or conversely (SL). It is essential to distinguish between EQ, LS, and SL as relationships between *chronometric values* (marked with index f) and as relationships between *perceived D values* (index φ). For instance, EQ_{φ}, i.e. perceived equality, very often occurs when the measured D values are in fact slightly LS_f or SL_f. In "triplets" or other groupings of three SE, the three D values often differ, one of them being the longest (L), one the shortest (S), and one in-between (designated O), e.g. LOS, SLO, or OSL. — The designations mentioned in this paragraph are not explicitly used in the RHYTHMSYVARD program but they are an important part of the conceptual framework and are often decisive for the Numericode labelling.

4. The Numericode labelling system

A necessary adjunct to the RHYTHMSYVARD program is a labelling system *defining the schematical metric position* of each D value from a series of raw data for a particular VERSION. This means that the investigator has to decide what SE should be considered to belong to a certain beat, a certain measure etc. For this purpose it is necessary to have either an existing notation or a simple "transcription" to start from. Of course, many calculations can be performed even without this type of classification and labelling, but a main feature of the RHYTHMSYVARD program is that such a classification should be possible. (Difficulties may occur, for instance, when a tune is played with several melodic ornaments and the like, and the investigator has to decide if a particular ornament starts "on-beat" or "off-beat".)

The labels used for this purpose have been borrowed from the Numericode system for thematic incipits (Bengtsson, 1967b). First the time signature is stated as in ordinary notation (C changed to 4/4 and ₵ to 2/2), followed by a double slash //, which is also used to mark the end of the MELODY/VERSION. The single slash / is used for bar lines. Time values are designated: 𝅝 = 1, 𝅗𝅥 = 2, ♩ = 4, ♪ = 8, ♪. = 6, 𝅘𝅥𝅯 = 3. Rests are marked by the letter R directly before the numeral, for instance, /𝅗𝅥 𝄽 ♩/ is coded /2 R4 4/.

If desirable or necessary, "dotting" may be represented by a dot, for example, 8. 6 meaning ♫ . As soon as such pairs of SE within a certain material are *clearly heard* to be different (e.g., often as LS but now and then as EQ), it is wise to keep them apart through the labelling, i.e. 8. 6 and 8. 8 respectively. This is particularly important when computing so-called differences (see 5.4 below), since the relationship between the D values in such pairs is then compared with a "mechanical" relationship, in the first case 3:1 and in the second 1:1.

Triplets and other *gruppetti* are labelled with the same numerals in brackets, for example, ♪♪♪ encoded as (8 8 8).

The whole procedure is easily understood by comparing the following Numericode labelling with the tune used as demonstration example (Fig. 1):

3/4 // 2 4 / 2 4 / 2 4 / 2 4 / 2 4 / 2 4 / 2 4 / 2. /
2 4 / 2 4 / 2 4 / 2 4 / 2 4 / 2 4 / 2 4 / 2. //

It is of the utmost importance that every measured D value in a VERSION is correctly paired with a proper Numericode label and that the total number of D values and labels is exactly the same. As already stated, the *final tone* in a VERSION often cannot be given a definite D value. In such cases this tone must therefore be disregarded. However, a corresponding "faked" D value may be included here as an approximation, if there are strong reasons for doing so and good "models" exist elsewhere within the material.

If a certain larger material of renderings contains several VARIANTS, difficulties might arise if they have to be labelled differently. This particular problem has not yet been tackled thoroughly, but as preliminary solutions may be mentioned (a) the use of different Numericode labellings for different variants, or

(b) treating a certain series of D values (by "grouping together" etc. at some places) *as if* a certain labelling were valid for more than one version.

5. Main tasks of the program

The main tasks of the RHYTHMSYVARD program are the following: (a) to make various types of *selections* (SEL) in a certain material of VERSIONS, (b) to make various types of *eliminations* (ELIM), e.g. of wrong D values, (c) to compute the sum of all D values, a mean value for the B duration and the average metronome value (M.M.) for every VERSION, (d) to undertake so-called *normalizing* (NORM) over a VERSION or parts of it, (e) to compute *deviations* or *differences* (DIFF) of D values from chronometrical and "mechanical" regularity, and (f) to compute D value *proportions* (PROP) at different metrical levels and in different ways.

5.1. Selections (SEL) and eliminations (ELIM)

In any material treated by the program, *selections* of groups of versions can be made, e.g. of all or several versions by the same performer or of versions that seem to have peculiar similarities, etc. This is made by the command "SEL:", followed by the identificators of all desired versions.

Another feature of the program is the possibility to perform *eliminations* by the order "ELIM", e.g.:

(a) Elimination of the *final tone in all versions,* ordered by "ELIM pS", where p means "position" and S stands for the Swedish word "slutton" (last tone).

(b) Elimination of D values in *certain metrical positions* (as defined in the Numericode labelling) *in all versions* within a TOTOBJ or a selection, ordered by the letter p followed by the designation for metrical position. Thus "ELIM pM8" means elimination of measure no. 8 in all versions.

(c) Elimination of particular D values *within a certain version,* ordered by the identificator followed by "ELIM" and either a position symbol or the number of the SE (one or more) to be eliminated.

The possibility of making eliminations should be used with care, and the criteria for making them should be stated. Arbitrary use of eliminations may introduce various forms of bias in the results and make comparisons with other results more complex.

5.2. Total duration, average B and M.M. values

By a subroutine, the following basic calculations are performed: (a) the *total duration* of a particular VERSION, computed as the sum of all measured and labelled D values (it is important to check that their total equals a number of full beats), (b) *the average D value of the beat* (average = arithmetic mean), and (c) the corresponding *metronome value.* In the demonstration example (Fig. 1), these three values are: (a) 17770 ms, (b) 395 ms, and (c) M.M. = 151.9.

The program also numbers all SE values and keeps track of their metrical

SE	75	43	75	41	76	40	77	42	74	45	75	42	73	43	137
M	118		116		116		119		119		117		116		137
2M	234				235				236				253		
NORM	42.2	24.2	42.2	23.1	42.8	22.5	43.3	23.6	41.6	25.3	42.2	23.6	41.1	24.2	77.1
	66.4		65.3		65.3		67.0		67.0		65.8		65.3		77.1
	131.7				132.3				132.8				142.4		
DIFF	-2.2	2.0	-2.2	0.9	-1.7	0.3	-1.1	1.4	-2.8	3.1	-2.2	1.4	-3.4	2.0	10.4
	-0.3		-1.4		-1.4		0.3		0.3		-0.8		-1.4		10.4
	-1.7				-1.1				-0.5				9.0		
PROP	63.6	36.4	64.7	35.3	65.5	34.5	64.7	35.3	62.2	37.8	64.1	35.9	62.9	37.1	100.0

SE	75	40	71	39	74	40	73	44	75	43	75	45	79	46
M	115		110		114		117		118		120		125	
2M	225				231				238					
NORM	42.2	22.5	40.0	22.0	41.6	22.5	41.1	24.8	42.2	24.2	42.2	25.3	44.5	25.9
	64.7		61.9		64.2		65.8		66.4		67.5		70.3	
	126.6				130.0				133.9					
DIFF	-2.2	0.3	-4.5	-0.3	-2.8	0.3	-3.4	2.5	-2.2	2.0	-2.2	3.1	0.0	3.7
	-2.0		-4.8		-2.5		-0.8		-0.3		0.9		3.7	
	-6.7				-3.3				0.6					
PROP	65.2	34.8	64.6	35.4	64.9	35.1	62.4	37.6	63.6	36.4	62.5	37.5	63.2	36.8

Fig. 1. The song *Elvira Madigan*, "Sorgeliga saker hända", notated in 3/4 time. All numerical values refer to a particular performance of this melody on the piano. The upper three rows give the measured Dii values in centiseconds for each sound event (SE, first row), each measure (M, second row), and each group of two measures (2M, third row). The final tone is omitted. The total duration of this version is 17770 ms, the average D value of the beat (corresponding to the quarter note) 395 ms, and the corresponding metronome value (M.M.) 151.9. The next three rows give the normalized values (for SE, M, and 2M, respectively), i.e. the Dii values expressed in per mille of the total duration. (According to a "mechanical" norm, a quarter note would correspond to 22.2‰ of the total duration, obtained by setting the total duration to 1000 and dividing it into the total number of beats, which is 45; 1000/45 = 22.2. A half note would thus correspond to 2×22.2 = 44.4‰.) The following three rows give the DIFF values (for SE, M, and 2M), i.e. the deviations from a "mechanical" norm expressed in per mille of the total duration. The DIFF values are also shown graphically in Fig. 2. The last row gives the relative proportion between the half note and the quarter note within each measure, expressed in per cent values. (According to a "mechanical" norm a half note would correspond to 66.7 % and a quarter note to 33.3%.)

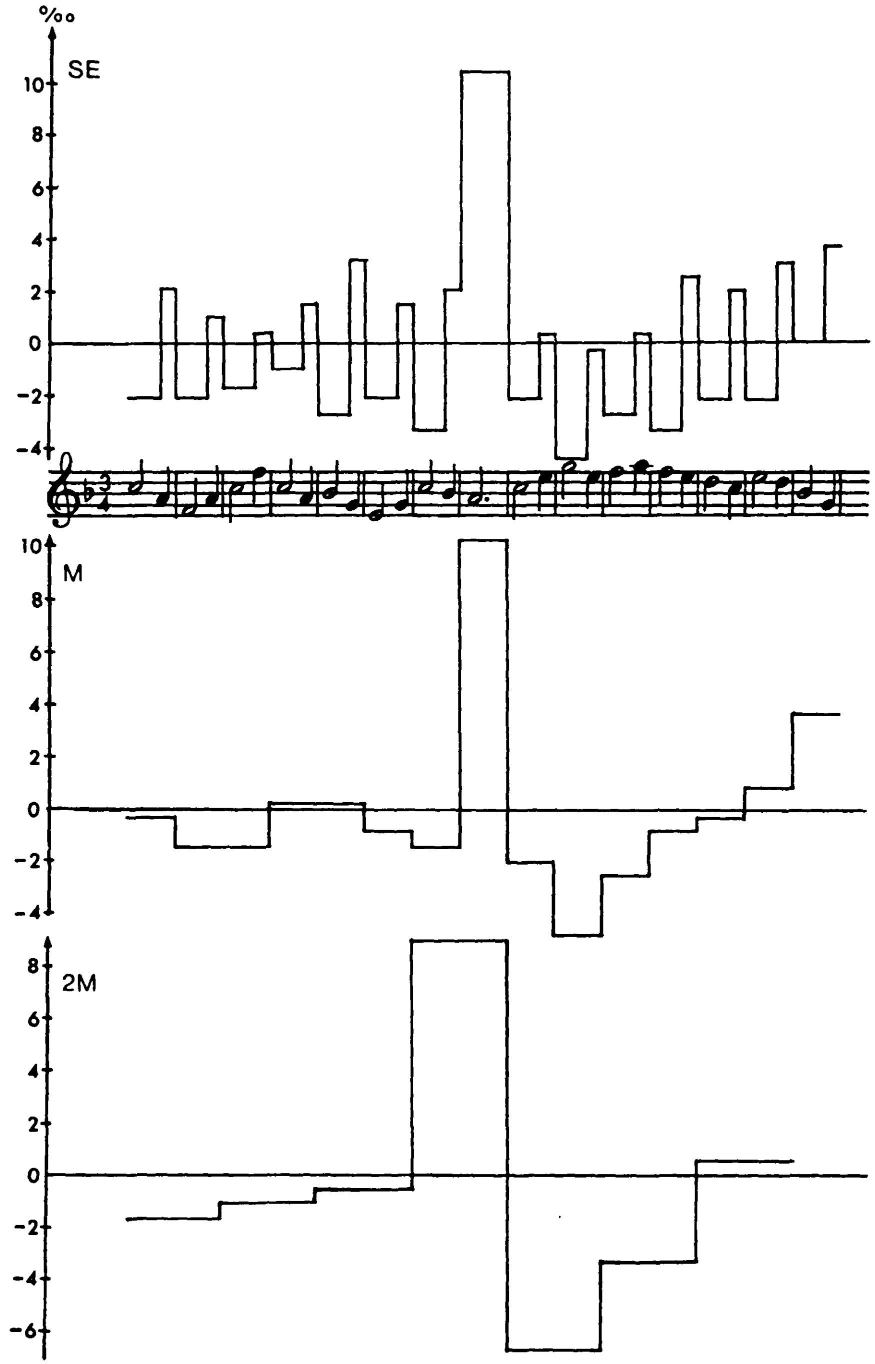

Fig. 2. Graphical representation of the DIFF values given in Fig. 1 for SE (upper), M (middle), and 2M (lower). The vertical axis represents deviations from a "mechanical" norm expressed in per mille of the total duration. Since the total duration is 17770 ms, one per mille unit corresponds to approximately 18 ms. The notation of the tune is added here to facilitate the reading and does not appear in the computer printout.

positions according to the Numericode labelling. It is recommended to number all SE in the notation (or transcription) that is used.

5.3. Normalizing (NORM)

In order to make the raw data from several versions comparable (for instance by eliminating overall tempo differences), this subroutine transforms all original measured D values to *fractions of a certain larger entity,* usually the total duration of the version. With such an entity put to 100 or 1000, all D values can be expressed as *per cent* or *per mille* of that entity. In the example (Fig. 1), the total D sum for SE 1—29 (the last note eliminated) is 17770 ms. Using "NORM over version", any particular D value d can be transformed into a per mille value p according to the formula p/1000 = d/D sum. The normalized values in the example appear in Fig. 1 below the D values. At present, these NORM values are not printed out by the program but are used for calculating the DIFF values according to 5.4.; however, the printing can easily be added if deemed necessary.

It is important to note that all tempo fluctuations as well as all kinds of deviations from "mechanical regularity" on any metrical level are retained if and only if the transformation is performed as "NORM over version". It would not be the case, for example, if "NORM over 4M-GR" was chosen. Then every such group of four measures would be set to the same unit (100 or 1000) and all differences between the D sums for the groups would be obliterated. "NORM over measure" is also possible, but the result would be equivalent to computing the proportions within the measure (compare 5.5.).

5.4. Differences (DIFF)

Any series of D values in a musical performance can easily be compared with a "rational-mechanical" norm derived from notation, i.e. notation used *as if* its signs for time values represented *exact* duration relationships such as 1:1, 2:1 etc. (e.g., as if all half notes were equal to 500 ms and all quarter notes to 250 ms). The measured D values can be treated as deviations or *differences from such a mechanical norm.* Suppose that two tones in a particular tempo should both be assigned the "mechanical" D value of 500 ms, but that their measured D values are 530 and 460 ms, respectively. The differences from the "mechanical" D value are thus +30 ms and —40 ms. Applying "NORM over version", assuming that the total duration (D sum) for the version is 12000 ms, these differences may be expressed as +2.5 ‰ (= 30/12000) and —3.3 ‰ (= 40/12000), respectively.

The DIFF values in the particular performance of the Swedish tune in Fig. 1 appear below the NORM values in the figure. In the DIFF values on the SE level an obvious "pattern" (that is, SYVAR-D) may be observed: the half notes are generally performed shorter and the quarter notes longer than what a "mechanical" norm would imply. Moreover, there is a lengthening of the single note in M8 (end of the first period), and a tendency towards a *ritardando* at the end of the tune.

To make this "pattern" easier to see, the DIFF values may be displayed graphically as in the upper part of Fig. 2, where the SE are plotted along the X-axis and the DIFF values along the Y-axis. The horizontal line at the zero DIFF value thus represents a truly "mechanical" performance. The real performance shows a very characteristic "pattern" around this zero line.

DIFF values and corresponding graphical representations can be generated by the program for all metrical levels that are of interest. In this case, besides the SE level already described, DIFF values and corresponding graphs are also given for the level of the notated measure (M) and the level of 2M-GR, as seen in Figs. 1 and 2. At these levels, too, there are characteristic deviations from what a "mechanical" performance would look like (= the zero line).

Plots may be produced by the program as in Fig. 2, i.e. a separate plot for each metrical level, or they may be combined into one single graph as if the three curves in Fig. 2 were plotted around a common zero line. Although the latter alternative saves space and some money, it is more difficult to read.

5.5. Proportions (PROP)

Any relationships between neighbouring D values, as well as between neighbouring positions within a version, can be calculated and compared to the mechanical norm. Such relationships are called *proportions.*

An LS relationship (for instance, 𝅗𝅥 ♩) with the measured D values of 640 and 360 ms may be expressed in several ways. The program allows for the following three methods:

(a) Percentage of their sum, i.e. 64.00:36.00 % (mechanical 66.67:33.33 %);

(b) The first value set to 1 and the second value expressed in relation to that, i.e. 1:0.56 (mechanical 1:0.50);

(c) Conversely, the second value set to 1 and the first value expressed in relation to that, i.e. 1.78:1 (mechanical 2:1).

Of course, proportions can also be calculated for groups of more than two tones, for instance for studying patterns such as /♩. ♪ 𝅗𝅥/ or /♩ ♫ ♫/. In such cases the percentage alternative is often preferable. Different PROP calculations may also be combined, as, for instance, in the second pattern (a) over the whole measure, (b) separately for 8. 6 on B2 and B3, and (c) for all 8. 6. cases.

In the last row of Fig. 1, percentage values are given for the relation between the half note and the quarter note within each measure. It is also possible to obtain the arithmetic mean (and standard deviation) of this relationship as an average over all measures. (This possibility also holds for the other two methods of expressing proportions.)

6. Some technical information

The program is written in FORTRAN IV. As applied on an IBM 370/155 computer it requires a memory storage of 156 K for most cases. Computer time for a case like the example in Fig. 1 amounts to one or two seconds.

The DIFF and/or PROP values resulting from the program are usually subjected to further analysis by statistical techniques such as analysis of variance, various correlation analyses, and factor analysis. The purposes and applications of such techniques will be described in following papers.

References

Bengtsson, I. (1967 a): On melody registration and 'MONA'. (Elektronische Datenverarbeitung in der Musikwissenschaft, ed. H. Heckmann, Gustav Bosse Verlag, Regensburg, pp. 136—144.)

Bengtsson, I. (1967b): Numericode — a code system for thematic incipits. (Svensk tidskrift för musikforskning, 49, pp. 5—40.)

Bengtsson, I. (1974): Empirische Rhythmusforschung in Uppsala. (Hamburger Jahrbuch für Musikwissenschaft, 1, pp. 195—220.)

Bengtsson, I., Gabrielsson, A. & Thorsén, S.-M. (1969): Empirisk rytmforskning. (Svensk tidskrift för musikforskning, 51, pp. 49—118.)

Bengtsson, I., Tove, P. A. & Thorsén, S.-M. (1972): Sound analysis equipment and rhythm research ideas at the Institute of Musicology in Uppsala. (Studia instrumentorum musicae popularis, ed. E. Stockmann, vol. 2, pp. 53—76.)

Bengtsson, I. & Gabrielsson, A. (1977): Rhythm research in Uppsala. (Music, room and acoustics, Royal Swedish Academy of Music, Stockholm, pp. 19—56.)

Castman, B., Larsson, L. E. & Bengtsson, I. (1974): Ett dataprogram (RHYTHMSYVARD) för bearbetning av rytmiska durationsvärden. Institutionen för musikvetenskap, Uppsala (mimeo).