Offprint from:

Nordic Prosody IV. Papers from a symposium
Edited by Kirsten Gregersen & Hans Basbøll

Odense 1987: Odense University Press

Pp. 105-114: Kjell Gustafson: A new method for displaying speech rhythm, with illustrations from some Nordic languages

Kjell Gustafson

A NEW METHOD FOR DISPLAYING SPEECH RHYTHM, WITH ILLUSTRATIONS FROM SOME NORDIC LANGUAGES

Introduction

Some important goals of speech rhythm research must be to establish the following:

1. What are the relevant timing units of speech?
2. What structures do these units constitute?
3. How can we describe, classify, and quantify the units themselves and the structures made up from them?
4. What is stress? Can phonetic stress be defined in an unambiguous way? What is the relationship between stress and rhythm?

Although a great deal of effort has gone into speech rhythm research, the above questions have not, on the whole, received satisfactory answers. One of the main reasons for this is that a suitable means of displaying graphically the relevant units, has been lacking.

Existing methods of displaying temporal units

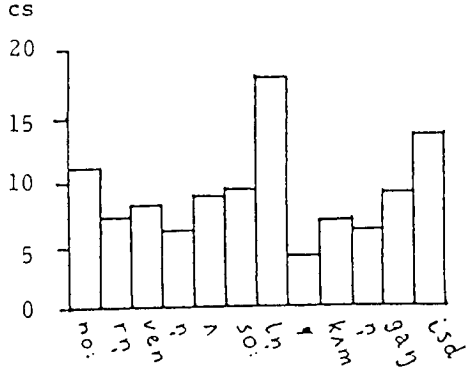
Fig. 1 shows the measurements from part of a Danish utterance displayed in three different ways that have been used in rhythm reasearch:

- Fig. 1a) displays the measured durations numerically (1); it is difficult for the brain to use this kind of information to make judgements about the rhythm of an utterance.
- In Fig. 1b) the measurements are displayed vertically, as histograms; this shows the durational relationship between the

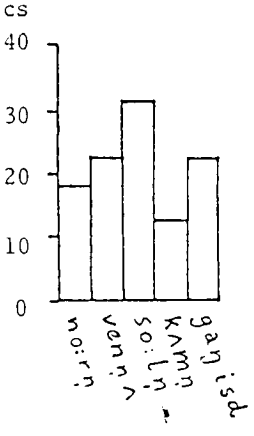
chosen units, but does not convey the relationship between each duration and true time.

cs	11	7	8	6	8.8	9.3	18	4	6.8	6	9	13.5	
	no:	rn	ven	n	ʌ	so:	ln	▲	kʌm	n	gaŋ	i	sdr ...
cs	18		22.8			31.3			12.8		22.5		

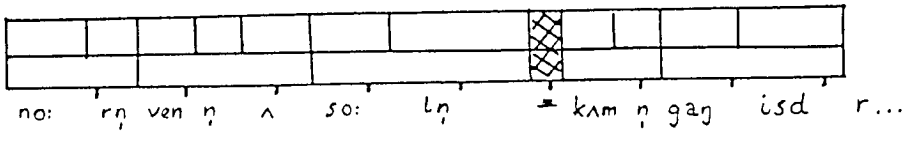
a) Numerical display



b) Histogram display: syllables



c) - groups of syllables



d) True time display; each calibration line represents 10 cs

Figure 1: Traditional methods of displaying temporal units and relationships; the symbol ▲ here and elsewhere indicates a pause

- In Fig. 1c) the syllables (and the pause) have been joined to groups; it takes a good deal of imagination to see how this display and that of Fig. 1b) relate to one another.
- Fig. 1d) displays the measurements along a line representing time; this gives a one-to-one relationship with true time, and it becomes possible to display simultaneously two different segmentations of the same utterance (as has been done here), but

it does not display very clearly the relative durations of the elements.

A new method

The purpose of this paper is to present a new method of displaying temporal elements graphically; I call it "Temporal Elements Displayed As Squares" (TEDAS). The method consists in simply displaying the measured durations both horizontally and vertically, thus combining the advantages of the histogram and the linear method illustrated above. In addition there are advantages that will be described below.

Fig. 2 shows the same measurements as in Fig. 1 displayed according to this method. This, I feel, gives an immediate impression

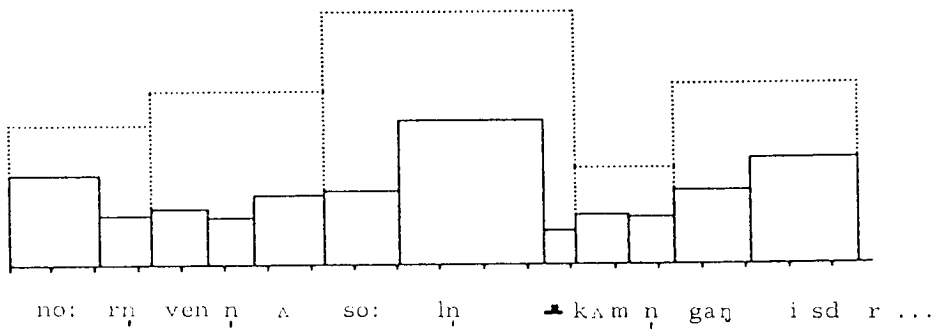


Figure 2: TEDAS display of Danish utterance: syllables and groups of syllables; calibration lines are 10 cs apart

of both the relationship between each syllable duration and between the individual durations and true time. I can only recommend that rhythm researchers try out the method for themselves. All that is needed is simple computer equipment and a simple computer program (I can supply a printout of mine on request) (2).

The justification for displaying the same measurements (those of duration) twice over, horizontally and vertically, is that the boxes display two different aspects of durations, 1) the dura-

tion in itself (the vertical dimension), and 2) the timing of the beginning and end of each element whose duration is displayed (the horizontal dimension); another way of looking at this is to say that the boxes display durations as a funtion of **true time.**

Advantages of TEDAS

1. With TEDAS one can superimpose the displays of different segmentations of the same utterance on one another, as was done in Fig. 3; this means that one can compare the duration of individual syllables with that of the group (foot) they belong to, or one can join together groups of syllables that appear to form a sub-group within the foot; in the diagrams shown here solid lines are used to display syllable durations and dotted lines are used for groups of syllables; it is also possible to use dotted lines to show for example the durations of individual segments within each syllable; this is a useful device for showing the effect of, for instance, differences in tempo on segment duration.

2. One gets an immediate impression of the temporal relationships of the utterance, and can study, simply by inspecting the diagrams, for instance the following:

1 - tempo and tempo changes; a consequence of this kind of display where the height of the boxes relates directly to the duration of the respective units is that one gets an immediate view of tempo and tempo changes: in the Nordic languages tempo mainly depends on the duration of the feet, which is reflected in the height of the relevant boxes;

2 - the degree to which the rhythm of an utterance is regular; thus this method will be invaluable to those investigating the questions of isochrony, stress-timing and syllable-timing;

3 - the interrelationship between tempo, durations of individual syllables, and durations of relevant groups of syllables; in this way one can see how certain patterns are typical of particular languages, or particular speakers or

speaking styles or rates, for instance with respect to the relationship between stressed and unstressed syllables.

This kind of display presupposes a segmentation of speech into discrete units; this must necessarily lead to some degree of arbitrariness - for instance, what segmentation corresponds most closely to the one the brain uses for the timing of syllable production and perception? I believe that this method will in the future help in determining the correct segmentation of syllables.

I have marked in the Appendix those syllables that have primary stress (accent) (according to my intuition) by diagonal lines in the respective boxes, and those that have secondary stress by little dots in the boxes; pauses have been marked by hatching. I have judged the presence of secondary stresses auditorily; it will be seen that in nearly every case a secondary stress coincides with what appears naturally to be the first syllable of a group; thus primarily and secondarily stressed syllables appear normally to have the characteristic of being group initial; I will hypothesize that the definition of phonetic stress is precisely this; this is one area which I intend to investigate myself and I believe that TEDAS will have an important role to play in future discussions of the nature of stress.

As it makes no difference in principle to use different scales along the horizontal and the vertical axes, I usually use a compressed version to save space. Some phoneticiáns have found this to be more in accordance with the perceptual image of the represented utterances.

To show the usefulness of the method I have measured the duration of syllables and pauses in the first part of altogether 10 recordings of 'The North Wind and the Sun' in Norwegian, Swedish, Danish, and Icelandic, and displayed them in the compressed mode; these diagrams are shown in the Appendix.

Space does not allow me to go further into details about the method or make detailed comments about the diagrams. I hope the

readers will be able to discern for themselves many interesting features of the rhythm of the utterances that are displayed there.

Conclusion

I believe the method described here will go a long way towards providing a tool for giving insightful answers to the questions set out at the beginning of this paper. The method is still slow to use, but this is due to the fact that measuring individual durations is very time consuming; when this process can be done automatically, TEDAS displays can become virtually simultaneous with the speech they portray. At that time one will be able to show instantaneous rhythm displays like the instantaneous pitch traces that can be produced on some pitch meters.

I would like to make one final claim: Isochrony in production and perception means equality of duration of the relevant timing units, subject to tempo changes; therefore, if adjacent boxes (those representing feet in the case of the Nordic languages) show a regular increase or decrease in size, this can be interpreted as indicating isochrony on this interpretation of the term.

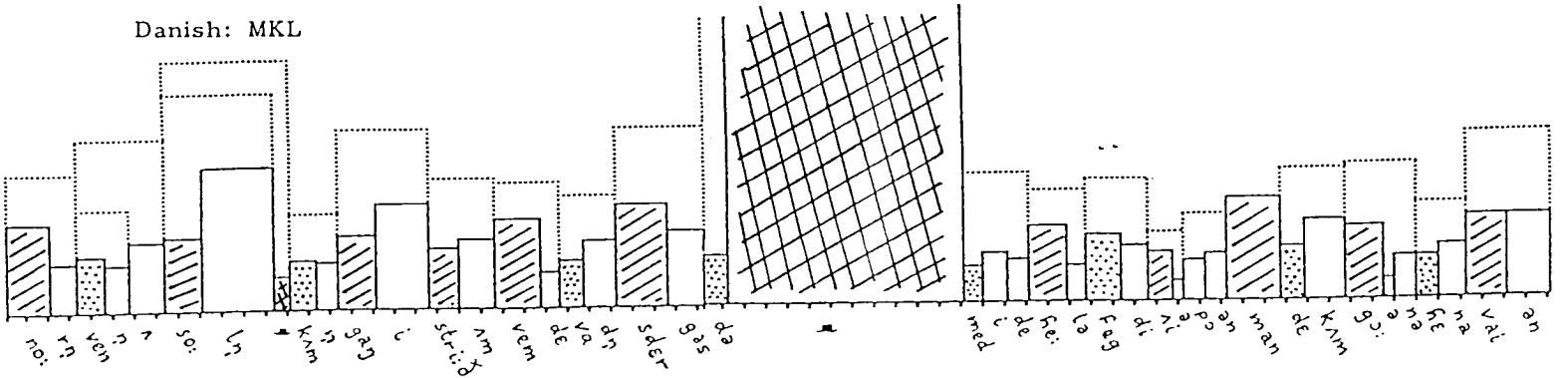
NOTES

1. The transcriptions used in this paper follow in the main those of the IPA booklet, but deviate from these when it was clear that the speaker had used a different pronunciation or a different wording. In the Danish transcriptions I have omitted to indicate the use of stød; this does not mean to say that I think that stød may not be of relevance to the timing of an utterance.

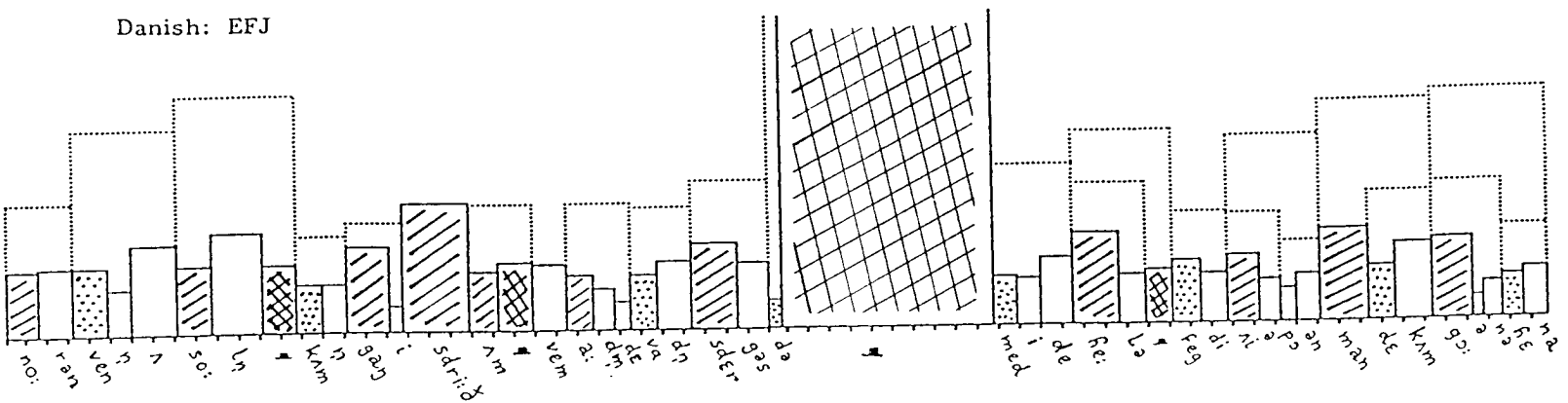
2. I am much indebted to Mr. Ceri Carlill, Oxford, who took the time and trouble to write this program for me.

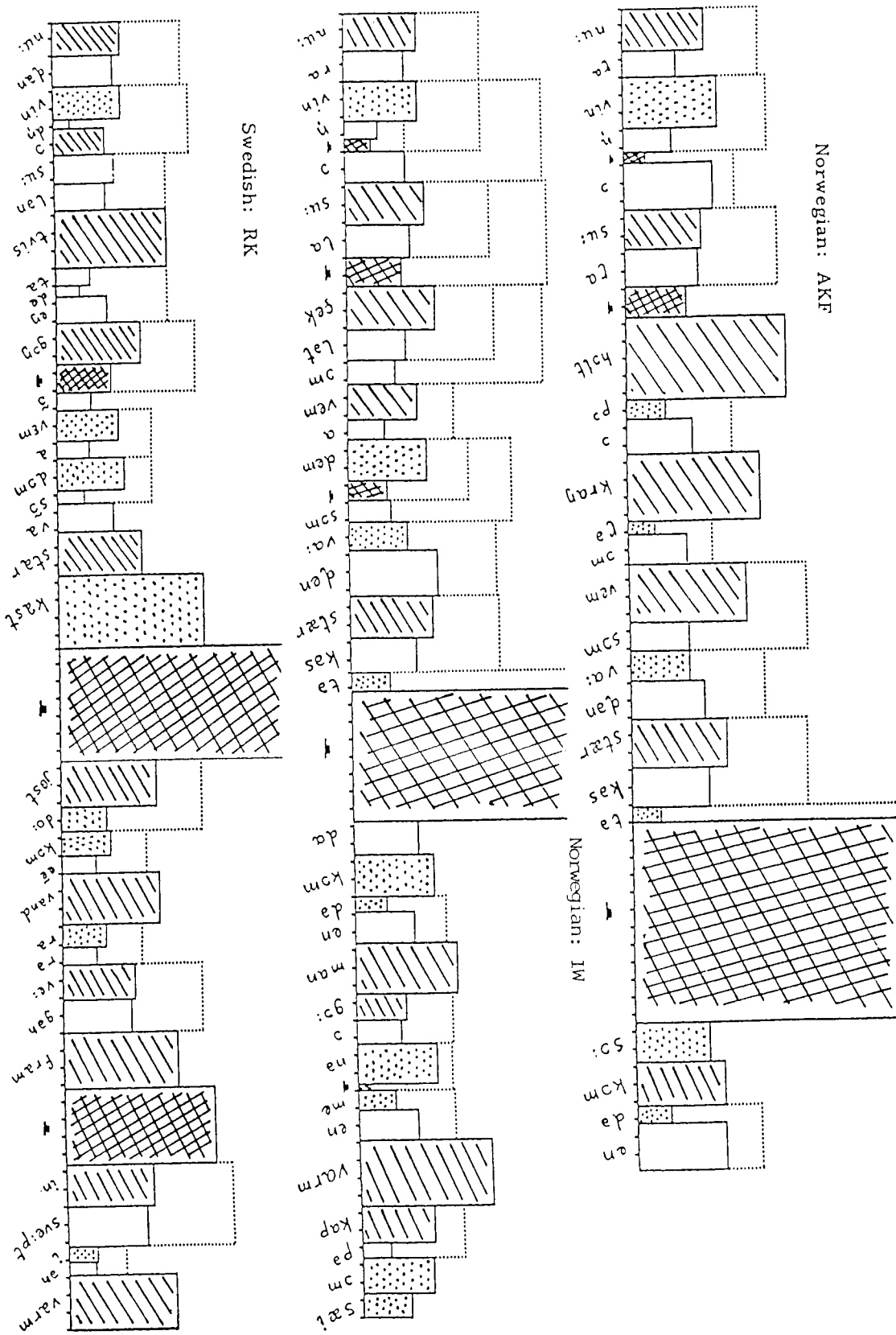
APPENDIX: TEDAS displays of Nordic languages

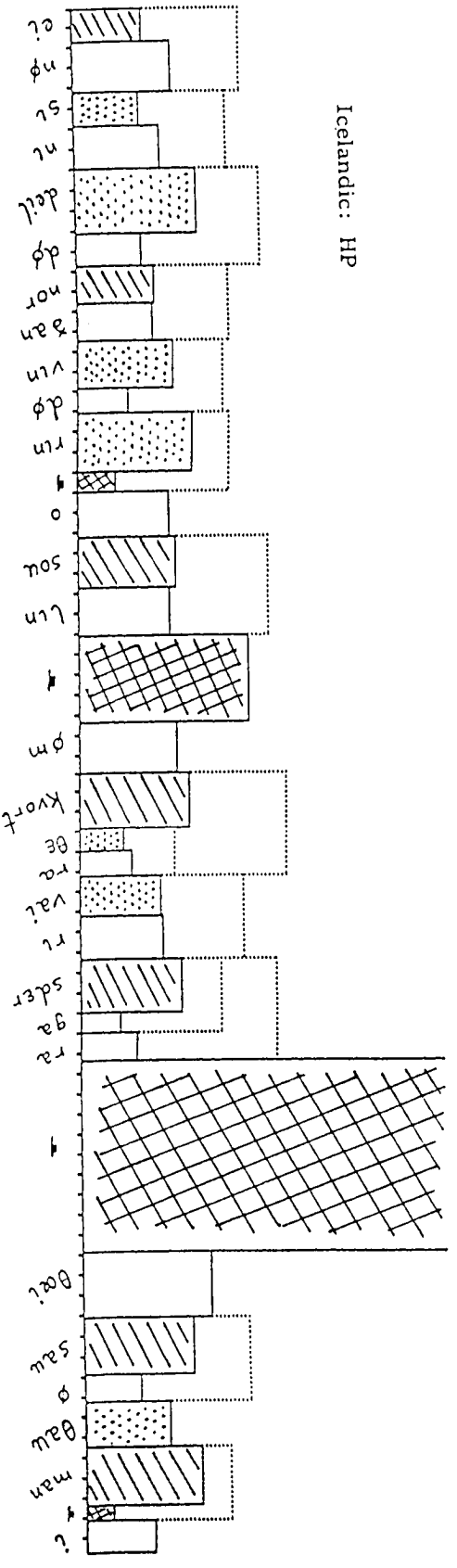
Danish: MKL



Danish: EFJ




Norwegian: AKF
Norwegian: IW
Swedish: RK


Icelandic: HP
Icelandic: JBH

Norwegian: KG (1)
Norwegian: KG (2)
Norwegian: KG (3)