Production and perception of phrases in some Nordic dialects

Eva Gårding and David House

Paper presented at the Sixth International Conference of Nordic and General Linguistics. August 18-22, 1986, Helsinki.

Introduction

Intonation helps us group words into phrases and phrases into sentences. This grouping is achieved with the aid of both connective and demarcative features. I have listed some of them here.

Connective signals may be:

(1) similarity of the elements of the group
(2) recurrent special patterns, rhythmic or melodic /1/
(3) tonal links between the elements (junctures)

Demarcative signals may be:

(1) breaks of similarity, for instance by the introduction of new elements like pauses, internal junctures, etc.
(2) special markers used to denote the beginning or end of a group.
(3) boundaries between recurrent patterns

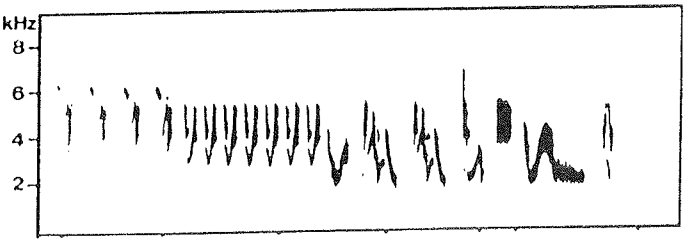
Phrasing occurs not only in speech but in all sorts of

time-ordered sequences, in bird song and music to give you two examples. Even here it is meaningful to speak of connective and demarcative signals /2/.

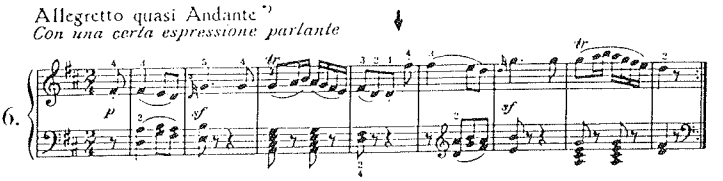
Take the song of the chaffinch for instance. It repeats the same pattern over and over again. As you can see in the spectrogram (Fig.1), the pattern has different groups connected by similarity of the elements, an introductory ti,ti,ti followed by a trill which changes to a slower trill which ends with a special marker, the concluding tiuiu. Note that the entire pattern is falling. The part of the song that you can see in the spectrogram is called a strophe by ornithologists (e.g. Bergmann & Helb 1982, from which the spectrogram is borrowed) and the groupings of similar elements are called phrases.

The next example is from Beethoven's bagatelle opus 33 No.6 (Fig. 1). In this line, up to the point which has been marked by an arrow, there are two groups of equal duration. The first group is falling with two markers, an introductory onset and a final strong long note. Beethoven marks this group with a slur. The second group starts with an onset and ends with a final marker consisting of three notes. This group is marked by three slurs, indicating three subgroups, one over the rising part and two over the final falling parts. The similarity with the intonation and rhythm of spoken sentences may be the reason why Beethoven added Con una certa espressione parlante to this piece. The grouping of music by melodic and rhythmic means is commonly referred to as phrasing. It is a very essential part of music and all important in its execution.

After this general introduction, I will only talk about phrasing in speech. Our talk is a report on work in



FRINGILLA CEOLEBS



BEETHOVEN OP. 33 NO. 6

Fig. 1 Examples of phrasing in bird song and music

progress. It aims at studying the production and perception of phrases in different prosodic systems. For today's talk we have chosen Swedish, Danish and Finnish /3/. I will talk about the production of phrases in a simple grouping experiment and make a summary comparison with real speech and David House will talk about some perception experiments.

Let me start with a rough definition of a prosodic phrase. A prosodic phrase is a part of an utterance in which accents or tones are coordinated in a unifying intonation pattern.

A grouping experiment

I will first describe a simple experiment in which speakers of different Swedish dialects and of Danish and Finnish were asked to say a series of fives, we called it a telephone number, as an ungrouped sequence and in groups of two and three and three and two, all in declarative intonation. The speakers used their respective languages and the Finns the colloquial monosyllabic form viis. With a sequence of numbers we meant to have stable experimental conditions with controlled influence from semantics and syntax.

You will now hear recordings of some of our informants and at the same time we will show you the corresponding pitch curves. The figures are practically self-explanatory. Just a few comments are needed. Look at Figure 2, for instance. The thick lines mark the intonation over the vocalic part of the syllable, the thin lines mark the voiced consonants and the dotted lines the voiceless ones. These dotted lines indicate the intonation movement masked by the voiceless segments. The parallel broken lines which enclose the curves are

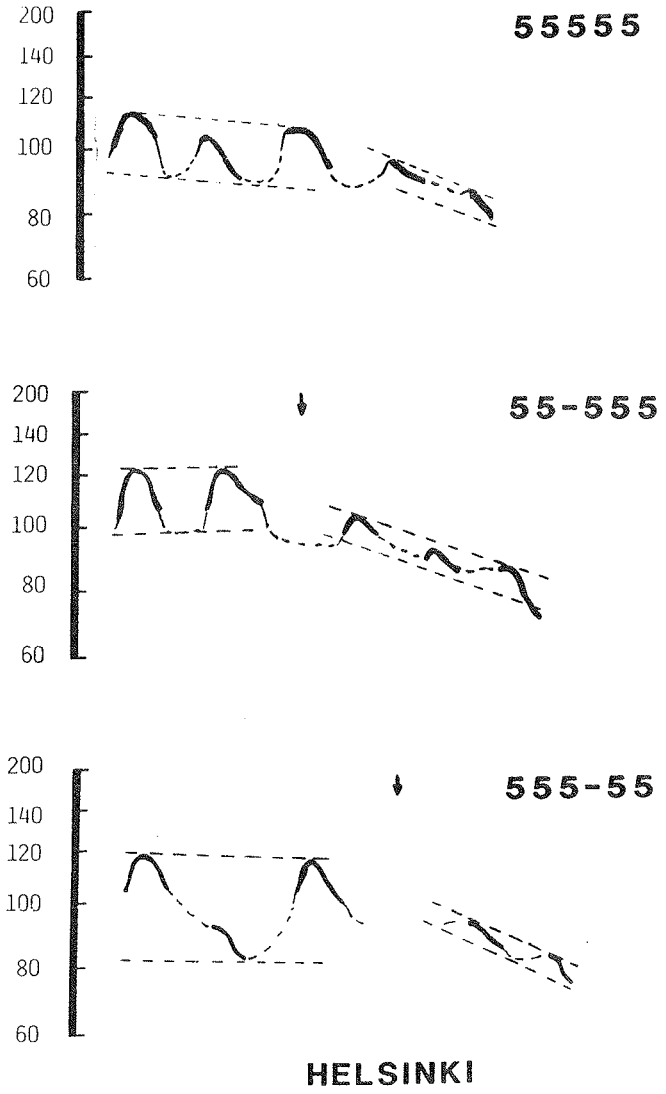


Fig. 2 Pitch curves from groupings of fives

auxiliary lines. We call them grids. They make it easy to see the general direction of the intonation and the point in time where there is a change of direction. This point is called a pivot and it is marked by an arrow. The width of the grid is correlated to stress. The wider the grid, the stronger the stress. A grid is convenient because it gives a combined measure of stress and pitch /4/.

(Here followed a demonstration of speakers from Helsinki, Stockholm, Gothenburg, Malmö and Copenhagen, accompanied by Figures 2,3,4,5 and 6.)

It is obvious from what you have just heard and seen that the speakers use both general and language-specific prosodic features to perform the required groupings. Let us look more closely at Figure 7 where I have combined the groupings performed by speakers from Helsinki and Stockholm.

The Finnish speaker has a wide, rather level grid for the first phrase and a narrow falling grid for the second. In the wide grid we see a special pattern in the group of three, which we can label High-Low-High. This gives it a trough-like tonal shape. The second group has a similarity pattern with three individual falls.

The Stockholm speaker has rising-falling grids to mark the two phrases. Like the Finn, she has a trough-like pattern in the first group but unlike him, she has a special marker, the rising movement to the high phrase accent for the non-terminal phrase and the same high phrase accent combined with a terminal fall to mark the end of the final group. The high phrase accent is typical of Central Swedish dialects (Bruce 1977). All in all, this is the tonal manifestation of an iambic or anapestic rhythmic group compared to the Finn's evenly spaced viises forming spondaic groups.

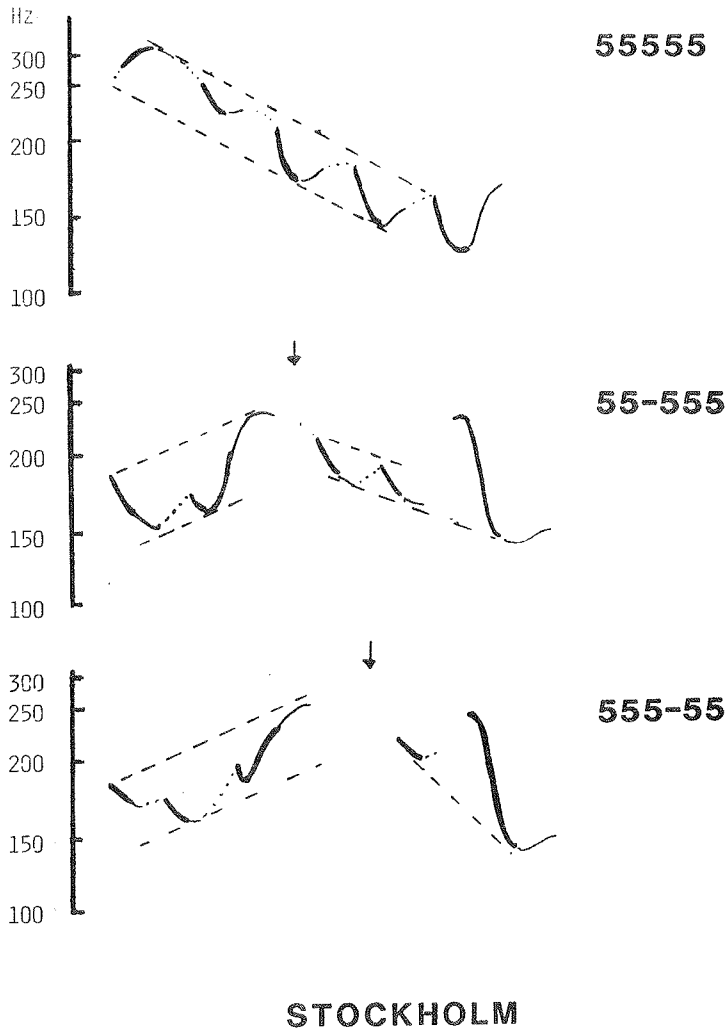


Fig. 3 Pitch curves from groupings of fives

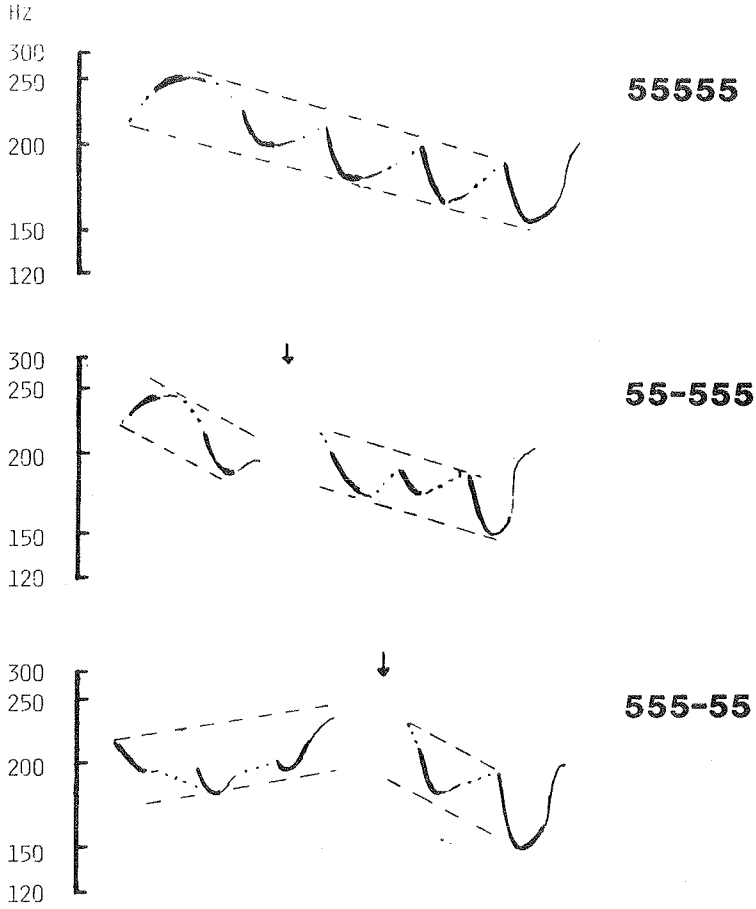


Fig 4 Pitch curves from groupings of fives

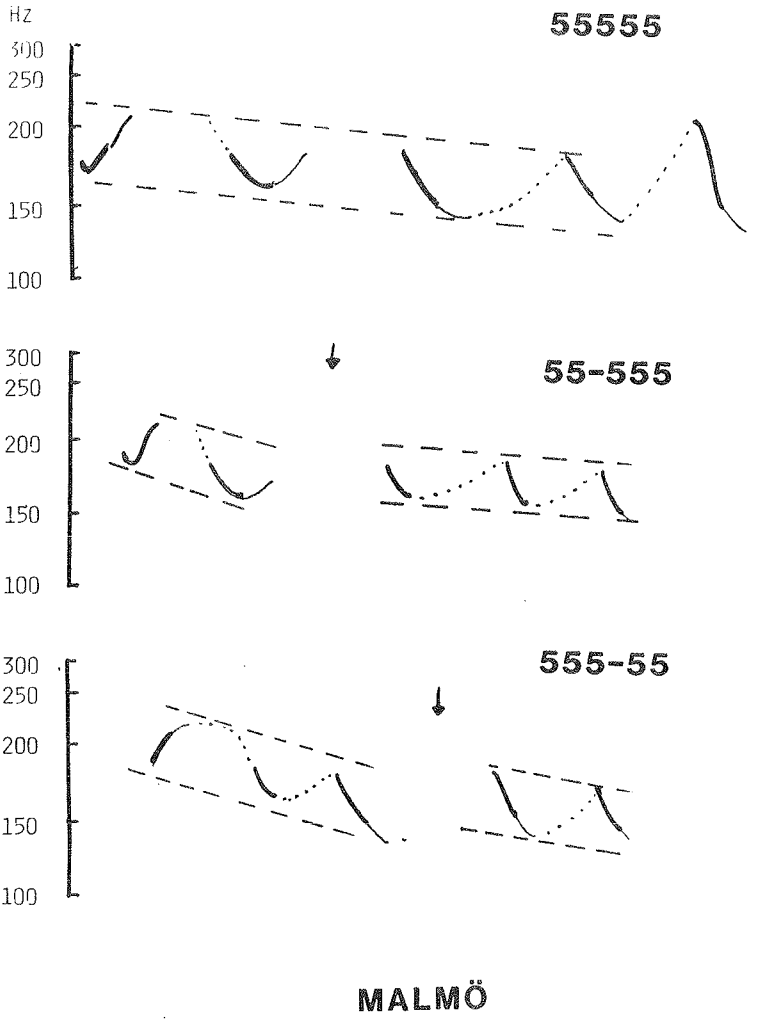


Fig. 5 Pitch curves from groupings of fives

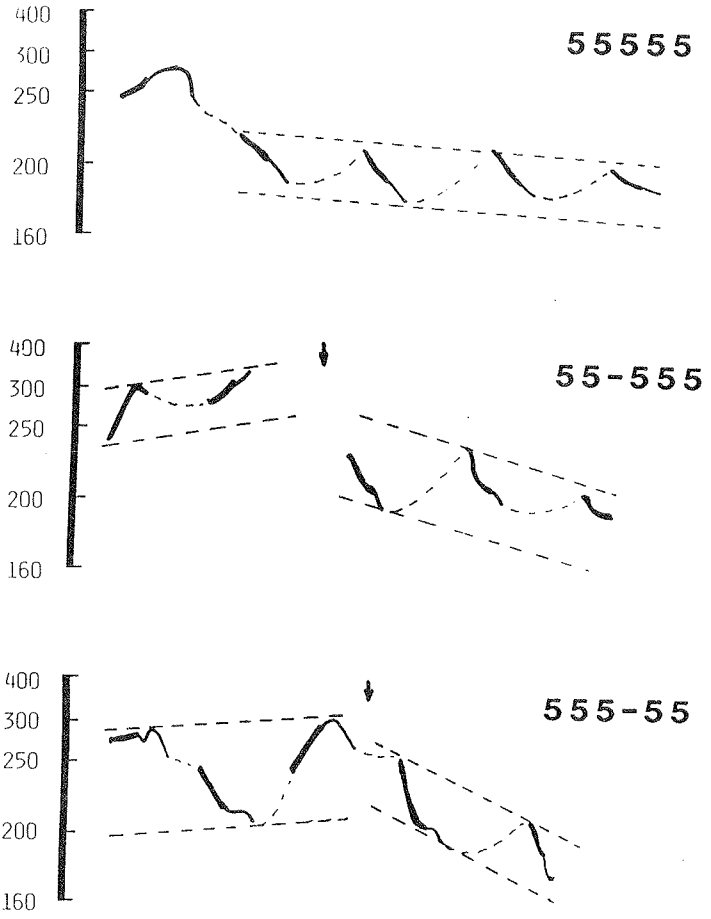


Fig. 6 Pitch curves from groupings of fives

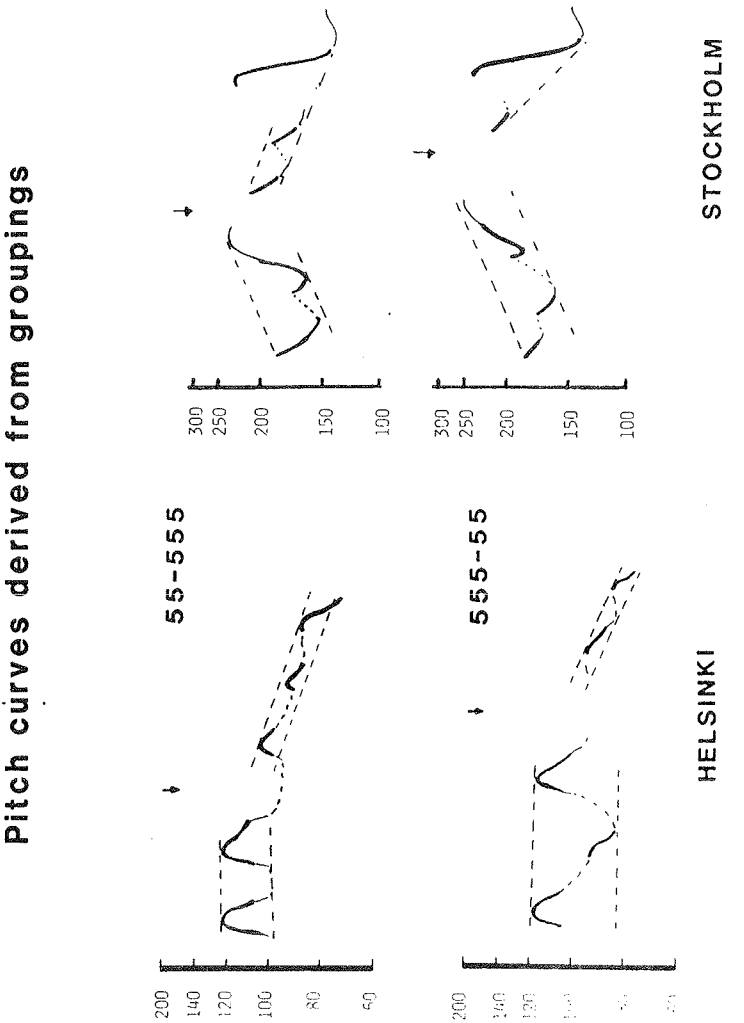


Fig. 7 Comparison of prosodic features used in groupings

These two speakers gave an example of a language-specific difference in grouping habits.

The Stockholm and Malmö speakers are examples of dialect-dependent variation (Fig. 8). The high phrase-final accent of Stockholm is in Skåne manifested as a low accent which is one of the reasons why these two dialects sound so different. Another reason is the Skåne speaker's spondaic rhythm which, not surprisingly, makes it more like Danish.

There are also examples of free variation in our material with some speakers using different strategies in their repeated renderings of the groups. Figure 4 gives an example of a Gothenburg speaker who wavers between two consecutive falls and a rise-fall for the groupings of the number.

Now let us look for some general principles present in all of the investigated dialects.

As demarcative signals all the speakers use pivots, manifested as pauses or a change of the grid range or grid direction. As connective signals all the speakers, irrespective of language and dialect, use falling similarity patterns in the second group, carrying declarative intonation, and special patterns in the first group (Fig. 9). The shape of the special pattern is determined by the terminal part of the group, a rise tends to be preceded by a fall and vice versa. The shape of the terminal part is dialect dependent. In groups of three, the mid accent is weakened and lowered which is a stable contribution to the pattern.

On the whole, this material reveals a very strict preplanning. When the last group is falling, the first group is level or rising. When the last accent of a group is

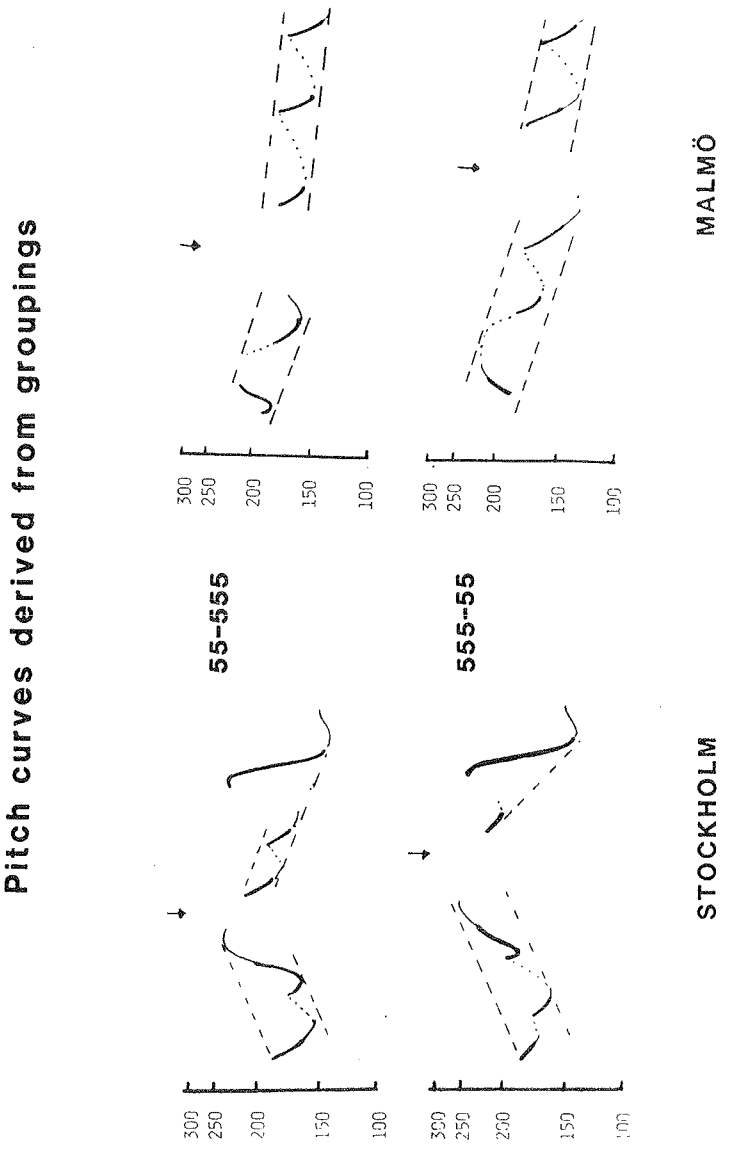
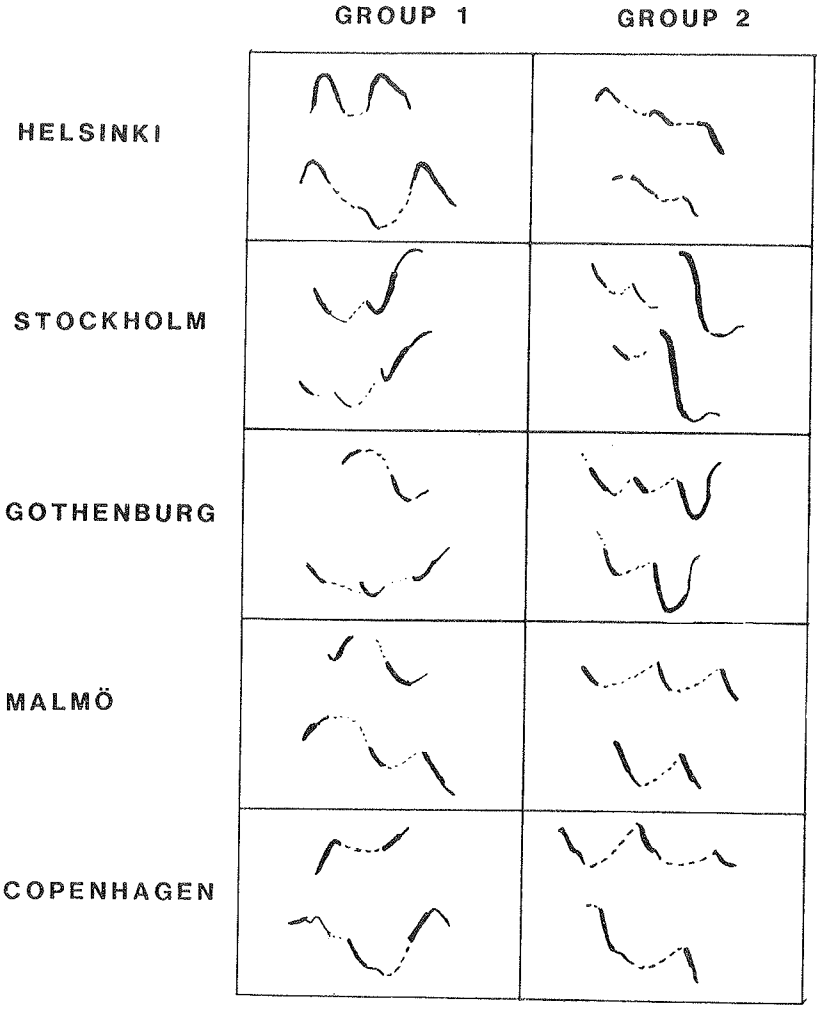


Fig. 8 Comparison of prosodic features used in groupings



Special patterns and similarity patterns

in groups of 55-555 and 555-55

Fig. 9 Special patterns and similarity patterns in groups of 55-555 and 555-55

rising, the preceding ones are falling, etc.

Phrasing in real speech

For this short talk we have decided to give preliminary answers to two questions. One has to do with connective signals, more precisely, the order-bound use of similarity patterns and special patterns. Is there any correspondence to this in real natural speech?

To find an answer to this question, we used a recording of a spontaneous conversation between Stockholm people, moderated by Bengt Loman in Malmö in 1966. (Gårding 1967a is an analysis of this material).

The answer is YES. Similar things do occur. Factors that determine the choice between special patterns and similarity patterns seem to be semantic importance and semantic coherence. In fluent narrative style the great majority of sentences seem to contain a focussed part, the rheme, which often comes early in the sentence. It is phonetically marked by a wide grid and its elements are closely connected in a special pattern. The rest of the sentence is backgrounded with similarity patterns in more or less compressed grids.

The other question that we asked of the material has to do with demarcation. What about phrase boundaries in real speech? Here one has to distinguish between precise boundaries and boundary regions. Precise boundaries are possible to detect in the acoustic record only when two accented syllables meet with an internal juncture between them /5/. In other cases, precise boundaries are replaced by boundary zones consisting of consonants or unaccented syllables of which some may be enclitic and others

proclitic. Their intonation forms a bridge between the adjoining accented syllables and this bridge is constructed according to the principle of the shortest way (Gårding and House 1985).

The implication is that in this case the precise boundaries cannot be detected by acoustic criteria, quite simply because they are not there. What listeners do, then, if they are asked to segment such a sentence into phonetic phrases, is to let syntactic and semantic criteria guide their segmentation (Gårding 1967a p. 51 ff.).

With this comment we have approached the perceptual part of our talk and David House will take over.

Perception of grouping

I am going to talk about some perception experiments which tested the relevance of the features we have observed. In particular we are interested in the relevance of the grid and the pivot as connective and demarcative signals /6/. For the tests 34 different synthetic stimuli consisting of sequences of fives were randomized and presented to 20 Swedish listeners according to standard procedure. Each stimulus received 100 responses. Listeners were asked to judge if the stimuli were grouped 2+3 or 3+2. As test stimuli we used manipulated variations of a natural Scanian fem 'five'. Note that pause is not one of the variables and that only tonal patterns are represented in the material.

Figure 10 shows a few examples of the stylized tonal contours used as stimuli. The results are given as percent "correct" response where "correct" means "expected" or "anticipated". For example, we expected stimulus 1, designed

Fig. 10 Examples of stylized tonal contours used as stimuli

to test a hat-like pattern as a connective signal to be grouped as 3+2 and, indeed, 74% of the responses were 3+2. Each stimulus has a counterpart, in this case stimulus 2 with the expected grouping 2+3. Here 72% of the responses were 2+3.

The next set of stimuli were designed to test the combination of a hat-like pattern and a large tonal movement as a demarcative signal. When both the connective and demarcative signals were used, the results improved considerably. Stimulus 3 now received 82% and stimulus 4 92% of the votes.

The following set shows the power of the special marker alone. Here we get 94% and 86%. More examples are rising and falling patterns and a trough-like pattern as connective signals. Finally, the trough-like pattern is combined with a special marker which again improves responses.

(Here the 12 sample stimuli were played.)

Our test shows that similarity of consecutive elements is a good connective signal. So are recurrent patterns (1,2 hats, 9,10 troughs, and 7 a rise and fall).

The special markers, the large-size tonal movements seem to be the best demarcative signals. It is interesting that these markers, both the falling and rising ones, are judged to be final elements. One possible explanation is that in Swedish a large final tonal movement is a characteristic feature of phrase stress or juncture. A break of direction or resetting of a grid is also a good demarcative signal. One exception is number 8 where this demarcative signal seems to be in conflict with a connective signal consisting of three very similar elements across the intended border.

The prosodic phrase in our model of intonation

Let me now come back to the definition of prosodic phrase. It is natural to apply the term prosodic phrase to each group of our two-group sentences. Our original definition of prosodic phrase in this talk is: a part of an utterance where accents and tones are coordinated in a unifying intonation pattern. We can now make the definition more precise: a prosodic phrase is a part of an utterance which is connected by a special pattern and bounded by pivots. A pivot can be a special marker or a change of direction or range of the grid. This is the definition used in our model of intonation/7/. It provides a useful tool for analyzing the aspects of intonation which help us group words into phrases and phrases into sentences, both in a Nordic and a general perspective.

SUMMARY

The grouping of time-ordered sequences uses both connective and demarcative features. Connective signals are similarity of the elements of the group and recurrent special patterns. Demarcative signals are e.g. breaks of similarity and special markers at the beginning or end of a group. The relevance of such signals for speech was first tested in a grouping experiment with numbers, 55555 contrasted with 55.555 and 555.55. The experiment was carried out with speakers representing Finnish (Helsinki), Swedish (Stockholm, Gothenburg and Malmö) and Danish (Copenhagen).

A common feature for all these dialects is that groups (phrases) are coordinated in a common unifying intonational

frame (tonal grid) and that discontinuities in this frame (pivots) correspond to phrase boundaries.

Another common feature in our material is the order-bound use of similarity patterns consisting of individual falls in the final group carrying declarative intonation, and the preference for a special pattern in the initial group. In initial groups of three, the mid accent is generally reduced.

The shape of the special pattern is dependent on the phrase-final accent which varies with dialect.

The number material suggests that there is a high degree of preplanning at two levels. At the sentence level the intonation of the last phrase seems to determine that of the first, and at the phrase level the shape of the last accent of the group determines the movements of the preceding ones. The governing principle seems to be one of contrast. A communication-carrying fall, as in a declarative sentence, is sharpened by a preceding rise and vice versa.

There is a certain amount of free variation for some of the speakers.

Features comparable to those found in the number material are observed in a recording of spontaneous Swedish speech. In a narrative style most sentences contain a focussed part marked by a wide grid and with its elements connected in a special pattern. The rest of the sentence is backgrounded with similarity patterns in compressed grids. As a rule, syntactic boundaries in a series of unaccented syllables are phonetically unmarked.

A complementary investigation of the perceptual relevance of connective and demarcative signals demonstrated that listeners can use the corresponding acoustic correlates,

grids and pivots, as cues to divide a series of numbers into two groups. Duration was not a variable in these experiments as the duration of each synthesized stimulus and the interval between them were always the same. The most powerful cue was a large tonal rise or fall which was judged to be a final element in a group.

NOTES

/1/.The special patterns are determined by phonological rules. In function and partly in form they are comparable to the results of Chinese sandhi rules.

/2/. In his classification of phonetic phenomena, Elert notes the importance of distinguishing connective features from demarcative ones (Elert 1970 p. 20).

/3/. Earlier reports are Gårding and House 1985 and House and Gårding 1986.

/4/. The notions of tonal grid and pivot with examples from several prosodic systems are explained in Gårding 1984 and Gårding 1985.

/5/. In Gårding's doctoral dissertation, internal juncture is defined as a marked syllable boundary in a phrase (Gårding 1967b p.33). Acoustic correlates are glottal stops or lengthening of the segments at the boundary.
The first comprehensive analysis of internal juncture had been made by Lehiste for English (1960).

/6/. That intonation by itself is a cue to grouping is already known from results reported in the literature (Collier & t'Hart 1975, Lehiste 1973).

/7/. There are at present different versions of this model. See e.g. Bannert 1984, Bruce 1977, 1982, Bruce & Gårding 1978, Gårding 1979, 1981, 1983.
An early presentation of the generative part of the model was given at the Second Conference of Nordic Languages and Modern Linguistics (Gårding 1973).

REFERENCES

Bannert R. 1984. Towards a model for German prosody. Working Papers 27, Department of Linguistcs, Lund University: 1-36

Bergmann H.-H.& H.-W. Helb. 1982. Stimmen der Vögel Europas. BLV Verlagsgesellschaft mb H. München

Bruce G. 1977. Swedish word accents in sentence perspective. Travaux de l'Institut de Linguistique de Lund XII. Gleerups, Lund

Bruce G. 1982. Experiments with the Swedish intonation model. In Preprints of Working group of intonation, Int. Congress of Linguistics, eds. H. Fujisaki & E. Gårding, Tokyo: 35-45 and Working Papers 22, Department of Linguistics, Lund University: 51-116

Bruce G. 1984. Aspects of Fo declination in Swedish. Working Papers 27, Department of Linguistics, Lund University: 51-64

Bruce G.& E. Gårding. A prosodic typology for Swedish dialects. In Nordic Prosody I, eds. E. Gårding, G. Bruce & R. Bannert. Travaux de l'Institut de Lund XIII: 219-228

Elert C.-C. 1970. Ljud och ord i svenskan. Almqvist & Wiksell. Stockholm

Gårding E. 1967a. Prosodiska drag i spontant och uppläst tal. In Svenskt talspråk, ed. G. Holm. Almqvist & Wiksell. Stockholm: 40-85

Gårding E. 1967b. Internal Juncture in Swedish. Travaux de l'Institut de Phonétique de Lund VI. Gleerup, Lund.

Gårding E. 1975. Toward a prosodic typology for Swedish dialects. In Proc. Nordic and Modern Linguistics II, ed. K.-H. Dahlstedt. Almqvist & Wiksell International, Stockholm: 466-475

Gårding E. 1978. Sentence intonation in Swedish. Phonetica 36: 207-215

Gårding E. 1981. Contrastive Prosody: A model and its application. Studia Linguistica 35: 146-166

Gårding E. 1983. A generative model of intonation. In Prosody: Models and measurements, eds. A. Cutler & R. Ladd, Springer: 11-25

Gårding E. 1984. Comparing intonation. Working Papers 27, Department of Linguistics, Lund University: 75-100

Gårding E. 1985. In defence of a phrase-based model of intonation. Working Papers 28, Department of Linguistics, Lund University: 1-18

Gårding E., A. Botinis & P. Touati. 1982. A comparative study of Swedish, Greek, and French intonation. Working Papers 22, Department of Linguistics, Lund University: 137-152

Gårding E. & D. House. 1985. Frasintonation, särskilt i svenska. In Svenskans beskrivning 15, eds. S. Allen et al. Göteborgs universitet, Göteborg: 205-221

House D. 1985. Sentence prosody and syntax in speech perception. Working Papers 28, Department of Linguistics, Lund University: 91-107

House D. & E. Gårding. 1986. Phrasing in some Scandinavian Dialects. Paper presented at the fourth Nordic Prosody

Conference in Middelfart, Denmark.
In preparation.

Lehiste I. 1960. An acoustic-phonetic study of internal open juncture. Phonetica 5. Suppl.

Lehiste I. 1973. Phonetic disambiguation of syntactic ambiguity. Glossa 7: 107-122

Loman B. 1966. A recording of a spontaneous conversation between Stockholm speakers. 2.5 hours on tape, in care of E. Gårding.