

**PHONETICS LABORATORY  
DEPARTMENT OF GENERAL LINGUISTICS  
LUND UNIVERSITY**



**WORKING  
PAPERS**

**15 · 1977**

MILAN BILÝ	Coreference rules described in terms of functional sentence perspective	1
EVA GÄRDING	Vergleichende Studien zur Prosodie schwedischer Dialekte	55
MAGNUS PÉTURSSON	L'aspiration des occlusives après [s]	67
HANS RANDLER	Identification test concerning isolated disyllabic Swedish accent 1 and accent 2 words	81
SIDNEY WOOD	A radiographic analysis of constriction locations for vowels	101
ROBERT BANNERT ANNE-CHRISTINE BREDVAD-JENSEN	Temporal organization of Swedish tonal accents: The effect of vowel duration in the Gotland dialect	133

Previous numbers and contents:

- WP 1 1969 KERSTIN HADDING Electromyographic study of lip activity  
MINORU HIRANO in Swedish CV:C and CVC: syllables  
TIMOTHY SMITH
- KURT JOHANSSON Försök avseende vokaltransitions riktning  
och dess betydelse för "plats"distinktionen  
bland tonande klusiler
- WP 2 1970 MONA LINDAU Prosodic problems in a generative  
phonology of Swedish
- WP 3 1970 GÖSTA BRUCE Diphthongization in the Malmö dialect
- EVA GÅRDING Word tones and larynx muscles
- KURT JOHANSSON Perceptual experiments with Swedish di-  
syllabic accent-1 and accent-2 words
- WP 4 1971 ROBERT BANNERT Hat das Deutsche zugrundeliegende  
stimmhafte Spiranten?
- EVA GÅRDING Laryngeal boundary signals
- KARIN KITZING Contrastive acoustic analysis of vowel  
phonemes, pronounced by some North German  
and South Swedish high school pupils  
(A summary)
- SIDNEY WOOL A spectrographic study of allophonic  
variation and vowel reduction in West  
Greenlandic Eskimo
- WP 5 1971 MICHAEL STUDDERT- Auditory and linguistic processes in the  
KENNEDY perception of intonation contours  
KERSTIN HADDING
- ANDERS LÖFQVIST Some observations on supraglottal air  
pressure
- WP 6 1972 ROBERT BANNERT Zur Stimmhaftigkeit und Quantität in einem  
bairischen Dialekt
- WP 7 1973 GÖSTA BRUCE Tonal accent rules for compound stressed  
words in the Malmö dialect
- EVA GÅRDING Constancy and variation in Swedish word  
PER LINDBLAD accent patterns
- WP 8 1973 EVA GÅRDING The Scandinavian word accents

- |            |   |  |
|------------|---|--|
| WP 9 1973  | PHONETIC<br>SYMPOSIUM   | Postgraduate students<br>Stockholm - Lund  |
|            | SIDNEY WOOD   | Speech tempo   |
| WP 10 1975 | ROBERT BANNERT<br>ANNE-CHRISTINE<br>BREDVAD-JENSEN                | Temporal organization of Swedish tonal<br>accents: The effect of vowel duration                        |
|            | GÖSTA BRUCE   | Stockholm accents in focus   |
|            | EVA GÅRDING<br>OSAMU FUJIMURA<br>HAJIME HIROSE<br>ZYUN'ICI SIMADA | Laryngeal control of Swedish word<br>accents   |
|            | KURT JOHANSSON  | Perceptual characteristics of vowels   |
|            | ANDERS LÖFQVIST   | Some phonetic correlates of emphatic<br>stress in Swedish  |
|            | BERTIL MALMBERG   | Niveaux, choix et systèmes approximatifs<br>dans le langage  |
|            | KERSTIN NAUCLÉR   | Some thoughts on reading and writing   |
|            | THORE PETTERSSON  | In favour of the archiphoneme  |
|            | EVA WIGFORSS  | Foreign accent and bilingualism  |
|            | SIDNEY WOOD   | What is the difference between English<br>and Swedish dental stops                                     |
| WP 11 1975 | MONA LINDAU   | Vowel feature  |
|            |   | A phonetic explanation to reduced vowel<br>harmony systems   |
|            | SIDNEY WOOD   | The weakness of the tongue-arching model<br>of vowel articulation                                      |
|            |   | Tense and lax vowels - degree of constriction<br>or pharyngeal volume?                                 |
| WP 12 1975 | ROBERT BANNERT  | The significance of vowel features in the<br>perception of complementary length in<br>Central Bavarian |
|            |   | Temporal organization and perception of vowel-<br>consonant sequences in Central Bavarian              |

- |            |   |  |
|------------|---|--|
|            | GÖSTA BRUCE   | Swedish accents in sentence perspective  |
|            | EVA GÅRDING   | The influence of tempo on rhythmic and tonal patterns in three Swedish dialects  |
|            | PETER KITZING<br>HANS-ERIK<br>RUNDQVIST<br>EWA TALO | Fundamental frequency of the voice in continuous speech. - Preliminary report on a device for determining mean and distribution of frequencies |
|            | ANDERS LÖFQVIST                                     | On the control of aspiration in Swedish  |
|            | MAGNÚS PÉTURSSON                                    | Etude glottographique de quelques consonnes islandaises  |
| WP 13 1976 | ANDERS LÖFQVIST                                     | Closure duration and aspiration for Swedish stops  |
|            |   | Oral air pressure in the production of Swedish stops   |
| WP 14 1976 | GÖRAN SONESSON                                      | Au sujet des fondements de l'analyse phrastique par Per Aage Brandt. Compte rendu  |
|            | THORE PETTERSSON                                    | Polsk fonologi i generativ tapping   |

Correspondence concerning this series should be addressed to:

Sidney Wood  
Phonetics Laboratory  
Lund University  
Kävlingevägen 20  
S-222 40 LUND  
Sweden



**PHONETICS LABORATORY  
DEPARTMENT OF GENERAL LINGUISTICS  
LUND UNIVERSITY**



**WORKING  
PAPERS**

**15 · 1977**

MILAN BÍLÝ	Coreference rules described in terms of functional sentence perspective	1
EVA GÄRDING	Vergleichende Studien zur Prosodie schwedischer Dialekte	55
MAGNÚS PÉTURSSON	L'aspiration des occlusives après [s]	67
HANS RANDLER	Identification test concerning isolated disyllabic Swedish accent 1 and accent 2 words	81
SIDNEY WOOD	A radiographic analysis of constriction locations for vowels	101
ROBERT BANNERT ANNE-CHRISTINE BREDVAD-JENSEN	Temporal organization of Swedish tonal accents: The effect of vowel duration in the Gotland dialect	133



COREFERENCE RULES DESCRIBED IN TERMS OF FUNCTIONAL SENTENCE  
PERSPECTIVE (FSP)

Milan Bily

Part I of this paper gives a survey of existing hypotheses (without referring, however, to such works as Grinder and Postal (1971), Brennan (1971), and Postal (1972b) which, in my opinion, just 'create' theoretical problems in order to use them as arguments in the controversy between interpretativists and supporters of generative semantics); Part II shows a new way to formulate constraints on pronominalization; Appendix A gives a short introduction to the theory of Functional Sentence Perspective, which is generally unknown in the West (except for some distorted fragments presented, for example, by Kuno (1972b) and Hinds (1975a)); in Appendix B some problems concerning English reflexives are discussed.

Part I - Earlier Studies of 'Pronominalization'

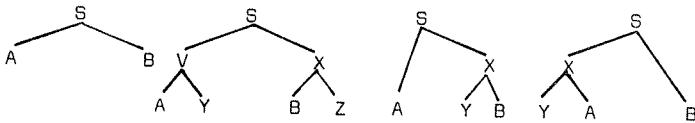
The standard transformational theory of pronouns and reflexives (Lees and Klima 1963, Chomsky 1965, Ross 1967, Langacker 1969) assumes that pronouns originate as fully specified NPs identical to their antecedents (and their 'postcedents') in deep structure. Transformations change these NPs into pronouns on the basis of morphological identity and intended coreference with other NPs. Thus a deep structure which can be roughly represented as "John<sub>3</sub> shaves John<sub>3</sub>" becomes "John shaves himself." (In Chomsky (1965) each NP has an associated index in deep structure and two NPs are coreferential if they have the same index.)



The following constraints on 'pronominalization' are generally accepted:

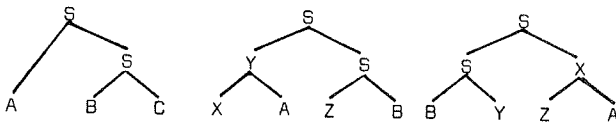
- I) NP<sup>a</sup> may not be used to pronominalize NP<sup>P</sup> if NP<sup>a</sup> and NP<sup>P</sup> are elements of separate conjoined structures and NP<sup>P</sup> precedes NP<sup>a</sup>.
- II) NP<sup>a</sup> may pronominalize NP<sup>P</sup> unless (1) NP<sup>P</sup> precedes NP<sup>a</sup>; and (2) NP<sup>P</sup> commands NP<sup>a</sup>.

(Langacker (1969, p. 167). Also Ross (1967) independently arrived at a practically identical condition on 'backward pronominalization'.) 'Command' is a technical term used to indicate the role of 'depth' in an embedded structure. Thus, the node A commands B, if a) neither A nor B dominates the other and b) the S-node that most immediately dominates A also dominates B. In the following structures A commands B and B commands A:



(The symbol 'S' always stands for a clause, the other symbols stand for any other part of an 'S' than 'S'.)

But in the following structures only A commands B, B does not command A:



Sentences (1) and (2) exemplify Constraint I; (3) - (6) exemplify Constraint II. (Asterisk-prefixed sentences are meant not to be able to express coreference between a pronoun and a NP, i.e. nothing is claimed

about their grammaticality. Nearly all example sentences are taken from the authors quoted in the bibliography. Either the original sentences are used or the sentence types - for example some sentences have been changed so that certain semantic abnormalities may be avoided for the sake of naturalness.)

- (1) Peter has a lot of talent and he should go far.
- (2) \*He has a lot of talent and Peter should go far.
- (3) John left town after he robbed the bank.
- (4) \*He left town after John robbed the bank.
- (5) After John robbed the bank, he left town.
- (6) After he robbed the bank, John left town.

However, Constraints I and II cannot explain many facts. There have been many attempts to mend these constraints. Lakoff (1968) showed that 'backward pronominalization' from a subordinate clause to a main clause (which is against Constraint II) is possible if the pronoun is not the subject of the main clause. Thus, it is possible to say (at least for some speakers):

(A: Why didn't Peter defend himself?) B: They silenced him every time Peter tried to speak.

Another (patently wrong) attempt was made by Kuno (1972a), using so-called 'direct discourse analysis', to explain sentences like (7) (with unstressed 'him'):

- (7) \*That John was the best boxer in the world was claimed by him.

The deep structure of (7) is said to be something which can be roughly described as 'John claimed: I am the best boxer in the world'. Since the subject of the embedded sentence is a pronoun from the beginning, there is no possibility to realize it as 'John' and there-

fore it is not possible to derive the ungrammatical (7). Unfortunately we can find sentences like (8) - (10), which should be equally wrong according to 'direct discourse analysis' but which are not. The fact that (7) is correct with stressed 'him' makes things even worse.

(8) That John was the best boxer in the world was never claimed by him.

(9) That John was the best boxer in the world was loudly and repeatedly claimed by him.

(10) That John was the best boxer in the world was claimed by him but nobody would believe such nonsense.

Postal (1972a, p. 48) postulates a so-called Global Constraint on Pronominalization in order to be able to make the difference between sentences like (11), (12) and (13), (14).

(11) Who killed his wife?

(12) It was Peter who killed his wife.

(13) \*Who did his wife kill?

(14) \*It was Peter who his wife killed.

"The Wh Constraint

Mark as ill-formed any derivation in which:

- a. there are two nominal constituents, A and B, in the input structure of a Wh Movement rule, where: (i) A is a pronoun, (ii) B is a wh form, (iii) A is to the left of B; and
- b. the corresponding constituents of A and B in the output structure of the Wh Movement rule, call them A' and B' respectively, are aligned such that B' is to the left of A'; and
- c. in the semantic representation, A and B (or, more precisely, their corresponding elements) are marked as stipulated coreferents."

However, there have been serious doubts as to whether there are any linguistic phenomena which need explanation in terms of a global constraint (e.g. Emonds 1973) and the global constraint on pronominalization in particular (Cole 1974). Besides, Cole (1974) shows that Postal's informants' judgements of his sentences are not at all fully representative for English speakers.

Another attempt was made by Postal (1971, p. 23) to save Constraint II with the following amendment: 'Backward pronominalization is banned across a copular verb of referential identity.' (emphasis in original). This restriction was motivated by the fact that there was no other way to prohibit 'pronominalization' in (15) while allowing it in (16):

(15) \*What annoyed him was my punching Bill.

(16) It was my punching him that annoyed Bill.

Bickerton (1975) believes that Postal's constraint is incorrect as it would predict noncoreferentiality for (17). However, Postal would probably claim that the copula in (17) is 'a predicate set of inclusion' (as he did for 'The man who wrote to her was a friend of Mary's.'). But then it remains to draw a clear borderline between these two possible meanings of a copula, which can be rather difficult, at least in certain cases.

(17) My punching him was Bill's major gripe.

There are more other 'bad cases of pronominalization' for which Constraint II does not suffice (cf., for example, Jackendoff 1972, Kuno 1972b etc.).

As for the above-mentioned Constraint I, it has been neglected by

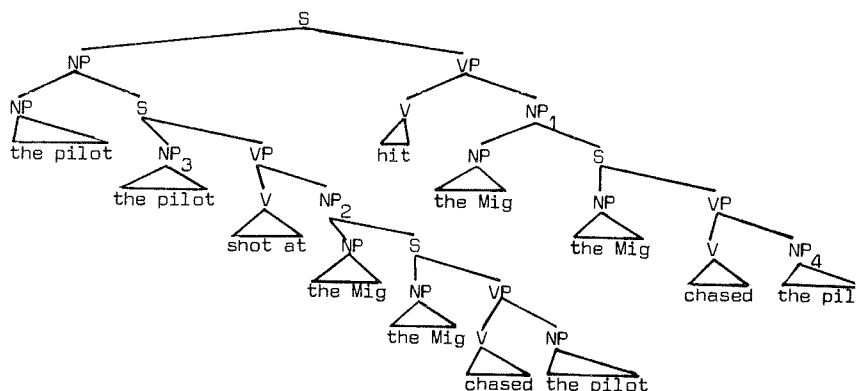
the majority of linguists. Only Postal (1971, p. 20) shows that one has to distinguish between "true coordination, where the conjuncts are, for instance, reversible without change of meaning, and pseudo-conjunction . . .". This restricts the operation of Constraint I radically, because even 'and' is very seldom a 'pure conjunct' without any causal, temporal or other implications affecting reversibility. Our sentences (1) and (2) cannot, then, be explained with Constraint I either. There is still another problem - there do exist structures of 'true coordination' where Constraint I should work but it does not:

(18) His wife and the woman Peter is living with have just met.

(18) is okay for many (or a majority of) speakers. Hinds (1975b, p. 332) tried to complete Constraint I with ". . . and NP<sup>P</sup> is stressed.", which would make the right prediction about (18) with stressed 'HIS' but it cannot explain the difference between (1) and (2) if 'he' is pronounced with reduced stress in both cases.

The most discussed sentence in papers on 'pronominalization' must be (19). This so-called Bach-Peters paradox (Bach 1970) is hard to explain with the pronominalization hypothesis. One can, of course, do as Karttunen (1971) did, i.e. one can suppose (in order to get rid of the infinite deep structure which would be necessary) that (19) has the following deep structure:

(19) The pilot who shot at it hit the Mig which chased him.



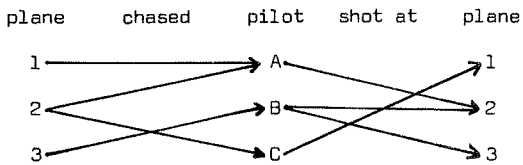
The  $NP_1$  pronominalizes  $NP_2$  and  $NP_3$  pronominalizes  $NP_4$ . But there are three possible meanings of this sentence, (Kuroda 1971) for which the transformationalists have to find three different deep structures. (According to the transformationalists every meaning of a given sentence must have a different deep structure.) Unfortunately for the transformationalists, it has been shown by Wasow (1973) that there is an algorithm for constructing an infinite number of deep structures for (19). Therefore transformationalists should discover infinitely many distinctive readings for (19)! The obvious impossibility of this means a hard blow to the transformational hypothesis of pronominalization.

Sentence (19) means either

- a) "The pilot who shot at the Mig that chased him hit it."
- or b) "The Mig that chased the pilot who shot at it was hit by him."
- or c) something which is neither a) nor b) and has the surface structure of (19). (Kuroda (1971) tried to explain this third possibility with reference to a hypothetical world where c) is possible without a) and/or b) being true, but in fact he himself proves, without real-

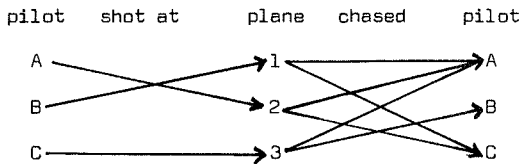
izing it, that c) is something else than a) and b). It is not necessary to refer to a non-existing 'data base'.) The fact that there do exist three meanings of (19) can be shown in the following 'data bases' taken from Kuroda (1971) and Karttunen (1971):

'Data base 1'



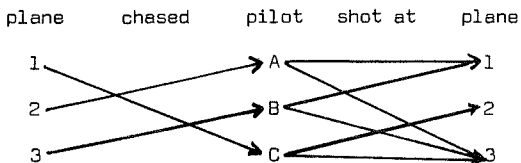
Only the pilot B and the plane 3 qualify for the meaning a), b) and c).

'Data base 2'



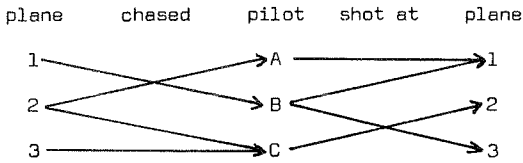
Only the pilot A and the plane 2 qualify for the meaning b) and c), but no pilot and plane qualifies for meaning a).

'Data base 3'



Only the pilot B and the plane 3 qualify for the meaning a) and c), but no plane and pilot qualifies for b).

'Data base 4'



The pilot B and the plane 1 qualify for both a) and c).

The pilot C and the plane 2 qualify for both b) and c).

As we can see, the meaning c) is sometimes 'synonymous' with a), sometimes with b), sometimes with both of them. a) and b) stand in a sort of hyponymical relation to c).

An attempt to save the pronominalization hypothesis was shown and criticized in Wasow (1975). He discusses the use of so-called bound variables (in sense of McCawley (1968, 1970)) which would make it possible to derive (19) from a deep structure which can be simplified as

$[x:\text{the pilot}], [y:\text{the Mig}], [x[\text{shot at } y]S_1 \text{ hit } y[y \text{ chased } x]S_2]S_3$ .

He shows that pronominalization in such a deep structure would generate certain ungrammatical sentences and it could not generate all grammatical sentences, because several syntactic rules in English are sensitive to the difference between pronouns and full NPs. If pronouns and their antecedents are derived from such variables, then these rules cannot apply correctly. Wasow gives, among other things, the following argument. If bound variables are accepted, the sentence with the simplified deep structure '[x:some burglars], [x shot a man who discovered [that x were in his house]S<sub>1</sub>]S<sub>2</sub>' will give 'Some burglars shot a man who discovered that they were in his house,' (via substitution of 'x' in accordance with Constraint II) or 'A man who discovered that some burglars were in his house was



shot by them.' Passivization must precede the substitution of 'x' to prevent generation of '\*They shot a man who discovered that some burglars were in his house.' As is known, 'there-insertion' in existential sentences is possible only when the NP in question is an indefinite one (\*There is the man at the door.). Wasow examines \*'Some burglars shot a man who discovered that there were they in his house.' and 'A man who discovered that there were some burglars in his house was shot by them.'. On the lowest transformation cycle in the latter sentence, there-insertion is applied, then passivization comes on the next cycle, and after that 'x' is substituted. But if passivization does not occur on the second cycle but only substitution, the former ungrammatical sentence is the unavoidable result; that is, the latter sentence can be generated only if the former ungrammatical sentence can.<sup>1</sup>

Wasow proposes an interpretative approach to the problem to save McCawley's bound variables. But this is already the method advocated by interpretativists (Dougherty 1969, Jackendoff 1968, 1972) who don't need to postulate such abstract deep structures like those proposed by Wasow. Interpretativists (led by the later works of Chomsky (Chomsky 1970, 1971 etc.) have returned to the position taken by Chomsky (1957). They do not postulate any 'fancy' abstract deep structures, their 'moderate' deep structures are syntactically motivated. They claim that

- a) Pronoun forms are inserted into deep structures like any other NPs.
- b) Coreferentiality between the full NPs and the pronominal NPs is a semantic judgement on the part of speakers/listeners and thus statements involving coreference shall not appear in the formulation of any transformation.

- c) Describing 'pronominalization' amounts to discovering rules of semantic interpretation of the type /NP<sub>a</sub>/ is  $\alpha$  coref /NP<sub>b</sub>, +pro/ just in case certain conditions are fulfilled .
- d) Deep structures with incorrect forms which cannot get any semantic interpretation are semantically ill-formed.

Among the arguments interpretativists use against transformationalists (besides the Bach - Peters paradox) are these:

- a) At least some pronouns must be generated by the base anyway (e.g. in 'She is beautiful.' combined with pointing at the person who was not talked about before).
- b) There are NPs like 'the bastard', 'the bum' etc. These 'pronominal epithets' can occur in certain environments where they function more or less as specialized pronouns. ('We asked Tom, but the bastard is too lazy to do anything.') There is no sensible way to describe, in a transformational framework, when a 'pronominalized NP' becomes a pronoun and when it becomes one epithet or another.
- c) Dougherty (1969) gives many other examples which are difficult to generate via a pronominalization transformation - e.g. 'Each of Mary's sons hated his brothers.' 'his brothers' cannot be derived from 'Each of Mary's sons' brothers' which would mean that every son hated all other sons and himself.<sup>2</sup>

(Dougherty has unnecessarily committed himself to a very strong claim: ". . . the set,  $\sum_A$ , of surface structure sentences which contain a proform that is understood anaphorically is a subset of the set  $\sum_N$  of surface structure sentences which contain a proform that is not understood anaphorically. . . . There is no a priori reason

why this should hold in English, but the fact that it does hold is linguistically significant." (Dougherty 1969, p. 511). There is, indeed, no reason why this should hold and in fact it does not – neither for reflexives, nor for such sentences where the semantics demands obligatory coreference ("Except for her laziness, Mary is an amiable person."). This has been criticized by McColl (1972), who is for the interpretative approach, and by Postal (1972b), who is against it. But Dougherty's statement just needs a slight reformulation to be correct. If we substitute 'syntactic surface structures' (i.e. an abstraction from the semantics of actual sentences) for Dougherty's 'surface structure sentences' and 'a proform that can be understood' for 'a proform that is understood', the claim is correct. Of course, the semantics of real sentences can determine that certain sentences get only the coreferent or noncoreferent reading.

Jackendoff (1972) expresses coreference in a "table of coreference". Each entry in the table consists of a pair of NPs and one of the relations coreferential or non-coreferential. Every possible pair of NPs in the sentence is included in the table. After the table is completed, it is subjected to so-called well-formedness conditions which determine whether it is consistent both internally and in relation to the rest of semantic representation. Jackendoff's rules for pronominal coreference are supposed to work as follows: the rules apply at the end of each transformational cycle and enter relations between pairs of NPs in the table of coreference. After the last cycle the noncoreferentiality rule, which says that every pair of NPs that have not been related by a rule of coreference will be marked as noncoreferential, ensures that every pair of NPs appear in

the table. Even Jackendoff's coreference rules contain a condition similar to Condition II on pronominalization, although he develops the concept of command from the earlier formulation to a more general one. For Jackendoff, the node A commands B, if a) neither A nor B dominates the other and b) the S-node or the NP-node (i.e. every node that defines a transformation cycle - according to Jackendoff, even NPs do that) that most immediately dominates A also dominates B. This makes it possible to give the right coreference interpretation to (20).

(20) His BROTHER visited John

There are also some other attempts to solve the problem of pronominalization with a combination of interpretative and transformational hypotheses, e.g. Harada - Saito (1972) where, by a mixture of the above-mentioned hypotheses, reflexives come about via transformations, while personal pronouns are generated in deep structure, or Teleman (1970), who proposes to derive pronouns and even other coreferent NPs from coreference-marked dummies in deep structure. These dummies get such surface structure representations as are able to express the intended coreference relations.

Generally speaking, the bulk of the works on pronominalization try to establish syntactic constraints on pronominalization. These basic syntactic constraints (the above-mentioned or some other - e.g. Culicover (1976), who uses the notion 'in construction with' instead of 'command') are then successively complemented (since they do not work) by new syntactic or other constraints. For example, a number of 'exception rules' are formulated for constituents bearing sentence stress. However, while the relations between sentence stress and 'pronominalization' are at times noted (Lakoff 1968, Postal 1971, Akmajian 1973,

Akmajian and Jackendoff 1970, Jackendoff 1972, Hinds 1975b), these amendments lie at the periphery of the study of coreference. They have never made an integrated coreference system; they are usually mere ad hoc patchworks.

One should not forget that the discussed rules of 'pronominalization' can be valid for a sort of 'normal language' only, i.e., a language in its 'basic functions', an abstraction from, for example, the metalinguistic function ('He says that Johnny wants a cake.' is okay as a compound of two languages - the baby language and the metalanguage of the adult speaker.), or the poetic function (A writer can certainly write: 'He thought that Peter behaved stupidly.' to indicate that Peter looked at himself from the outside, as a neutral observer.) Restrictions on 'pronominalization' are also valid only for a sort of coreferentiality intended by the speaker. The sentence 'But it was HE who said that Peter was not here.' is okay in this situation: Peter, who I did not know, pretended he was somebody else and said that Peter was absent. I told it later to a friend of mine who knew Peter and the friend pointed Peter out. I objected: But it was HE etc., by which I showed my reluctance to accept that Peter and the person I saw were identical. Then 'HE' and 'Peter' have the same referent, but I do not know it/ refuse to believe it.

Part II - Coreference Rules Described in Terms of  
Functional Sentence Perspective

In this attempt to show a new means of explaining the problem of coreference, the interpretative theory is combined with the theory of Functional Sentence Perspective (FSP). The aim of this paper is not to present a complete theory of pronominal coreference, e.g. the

question of coordinate sentences and coordinate NPs is avoided and English reflexives are just mentioned in Appendix B as they represent a language specific problem (for example, see Spangler (1970), while 'pronominalization' works practically identically in many (perhaps all) languages, however, an attempt is made to show that it is possible to do without the generally accepted 'precede and command' rule. For those who may be unaccustomed with the theory of FSP, as it has been developed by the Prague school during the last forty years, Appendix A is included in this paper. This paper's bibliography also contains several papers by Firbas, Daneš, Svoboda, Pala, and Dvořáková concerning FSP.

Although FSP belongs to text linguistics, it is possible to speak about the distribution of Communicative Dynamism (CD) even in an isolated sentence which we analyse out of its context. When no context is known, everybody 'reconstructs' a type of context in which the sentence in question could be used. It is always the 'normal' use of an isolated sentence, i.e. the sentence taken at its 'face value' is the sentence uttered with the 'normal', 'unmarked' intonation. That's why only the 'marked' stresses in example sentences are indicated by capital letters, otherwise all sentences are supposed to have 'unmarked' sentence stress. Two otherwise identical sentences with different sentence stresses (i.e. with different theme - rheme structures) are NOT identical sentences. Now and then, generativists realize that certain stresses change coreference possibilities; generally they ignore these differences and analyze the ambiguous written sentences only (i.e. they behave as if the written language were prior to - or, at least, equal with - the spoken language; in fact the written language is a simplification of the spoken one), the consequence of which is that

they may say whether coreference is possible (for at least one way of reading the written sentence), but they have no chance to discover why.

It is not always easy to analyze the gamut of CD in a given sentence. But the intuition and reasoning of linguists can be checked in several objective ways. One of these is a careful translation into a language in which FSP is the leading principle of word-order, e.g. into one of the Slavic languages. The rheme proper (the most rhematic word or phrase) bears also sentence stress. Another method of discovering theme-rheme structure is the method of questions. One can construct series of questions which the given sentence (with the given intonation) suits as an answer. (Except for such general questions like 'What has happened?'.) The element(s) which must be present in all questions belong to the theme of the analysed sentence and the element which is absent in all the possible questions is the rheme proper.

There have been several attempts to incorporate certain notions as 'theme', 'rheme' (or 'old information', 'new information', or 'pre-supposition', 'assertion') in the rules of 'pronominalization' (Kuno (1972b) and its enlarged version Kuno (1975), Hinds (1975a), and above all Bickerton (1975) who has completely discarded the old syntactic framework). However, these attempts have not been very successful because their theoretical equipment is insufficient for this aim. Hinds and Kuno (who both retain the old discredited syntactic framework) try to use the theory of FSP in their analyses, but their knowledge of Functional Sentence Perspective is minimal and distorted. For example, Hinds believes that every 'by-phrase' in passive sentences

must be rhematic because of its final or 'near final' position! The fact that most 'agent-nodes' are deleted (even when the agent is known) and that the agent can be expressed by an unstressed pronoun (i.e. by a thematic element) shows the fallacy of such purely mechanical judgement based on sentence linearity only. (Hinds is not alone, however. Since the American generativists started thinking about 'pre-suppositions' and similar notions, they have produced a lot of incredible statements the fallacy of which is obvious to everybody who knows the theory of FSP. One has to regret that these works begin where the Prague school was about forty years ago instead of using the accumulated results of Czechoslovak linguists.) Of course, nobody can stop Hinds if he wants to use some terms in a way defined by him, but Hinds believes that he applies the Prague theory of FSP, which he praises and makes propaganda for!

One of Hinds' most astonishing statements is "A function (i.e. purpose) of pronominalization is to indicate that the referent [!?] of the pronoun is considered thematic material." (Hinds 1975a, p. 91). According to Hinds, 'things in our world' are thematic or rhematic! Even Kuno (1975, p. 280) seems to believe in existence of 'predictable' and 'non-predictable' referents. There may live some 'thematic birds' or 'predictable butterflies' in the USA and Japan but they have never been seen in Central Europe. Kuno's lack of insight into FSP causes the terms he uses to be both numerous (theme, contrastive listing, exhaustive listing, neutral description, predictable theme, unpredictable theme, contrastive theme, known part of sentence, unknown part of sentence, old information, new information etc.) and extremely vaguely defined. ("I can only say that the theme is what the rest of sentence is about." - Kuno 1975, p. 277.) Of course, even the Czecho-



slovak linguists use at times such expressions when trying to popularize the concepts of FSP but one should (e.g. as Firbas has done) make some exact definitions, too. (In FSP terminology it is sufficient to speak about degrees of Communicative Dynamism (CD) as the other terms are used for the sake of convenience only:

'theme' = element(s) carrying a low degree of CD

'theme proper' = element(s) carrying the lowest degree of CD

'rheme' = element(s) carrying a high degree of CD

'rheme proper' = an element carrying the highest degree of CD  
etc.)

According to Kuno's definitions, it would be possible to claim that 'John' in 'Look, John is coming here.' is the 'theme' (more exactly: 'the unpredictable theme') as 'John' is what the rest of sentence is about. But Kuno suddenly says that this sentence is themeless since "there has been no previous mention of him [= John] or no expectation of his coming . . ." (Kuno 1975, p. 278). On the other hand Kuno (1972b, p. 308) says that "unpredictable themes appear when new topics are introduced". But this would be exactly the case of the above-mentioned sentence! Besides, Kuno (1975, p. 277) states that "The theme must be discourse anaphoric . . ." but it does not have to be "old predictable information". The first part of this statement clashes with the quotation about 'unpredictable themes'. What is meant by 'predictable themes' is mysterious, too. The fact that 'Tom' in the following dialogue is called 'unpredictable contrastive theme' is another example of Kuno's elastic and contradictory definitions:

"Speaker A: I understand that John, Bill, and Tom all teach high school.

Speaker B: John does, and Bill does, too, but Tom does not." (Kuno 1975, p. 277). 'Tom' is as much (or as little) 'predictable' in this

context as 'He' is in the dialogue on the same page about which Kuno claims that it is a 'predictable theme'!

"Speaker A: What does John like? Speaker B: He likes fish."

Both 'Tom' and 'He' are 'predictable' in the same way: If these words were made unidentifiable when the sentence they belong to were said, one could guess what should have been there. (The criterion is Kuno's own.) Otherwise they are not predictable at all. Cf. the possible answers of Speaker B: 'John does, and Bill does, too, but who is the third guy?', 'Why are you asking me?'. In fact, it is even worse. It seems impossible even for Kuno to remember what he means with his vague claims. In his latest version (Kuno 1976), he repeats that 'He likes' of the above written sentence 'represents old, predictable information because even if that part of sentence is garbled, it is recoverable from the preceding context'. Then he claims again that 'John', 'Bill', and 'Tom' represent 'unpredictable information' according to the same criterion'. (Kuno 1976, p. 120-121). But suddenly (Kuno 1976, p. 181) we read:

"Speaker A: Who do you like better, John or Mary?

Speaker B: Between these two, I like him better."

". . . the subject of like is him (= John), which is unstressed because it does not represent unpredictable information" (emphasis by the author of this paper). If we use Kuno's criterion again (" . . . if that part of sentence is garbled . . ." etc.) we shall find that 'him' does represent unpredictable information - the answer could have been " . . . I like her better".

It seems to me - as far as it is possible to say something about a language one has only a 'second-hand knowledge' about - that the fa-

mous Japanese particles are multifunctional and that's why their total identification with a sort of FSP markers (which is probably what Kuno does) causes such a chaos. Kuno also uses some examples of very dubious value:

"Speaker A: Who killed Mary?

Speaker B: \*With an accomplice that John hired, he killed Mary." (Kuno 1976, p. 123). The sentence is, of course, at least very strange but it has nothing to do with pronominalization. The following versions would not be better:

Speaker B: ?? With an accomplice that he hired, John killed Mary.  
or Speaker B: ?? With an accomplice that Jane hired, John killed Mary.

All these answers do not suit as answers to the question, irrespective of where the pronoun is placed or not.

Kundo (1975) presents ten rules of pronominalization (one of them is the traditional Constraint II), which are largely unrelated and whose validity is impossible to test because of his vague terminology and numerous definitions and 'counter-definitions'. Some of the rules are quite mysterious and remain unexplained (e.g., "If a given noun phrase in a sentence has a discourse anaphora, pronominalize all but one occurrence of that noun phrase in the sentence." - Kuno (1975, p. 280).), others are unnecessary (e.g., "Do not pronominalize the noun phrase of the exhaustive listing interpretation." - Kuno (1975, p. 280) - which is completed (p. 281) with the possibility to have stressed pronouns of 'exhaustive listing'. As Kuno's 'exhaustive listing' translated into a more general FSP term is 'rheme proper', it is obvious that it cannot be an unstressed pronoun and there is no need to formulate a special rule of 'exhaustive listing').

The only statement of Kuno's which is correct (after having been "translated from 'Kunoese' into English" is: All pronouns must have some ante cedents, either mentioned before or 'given' in the situation; i.e. so-called backward pronominalization is possible only when the referent is already 'known'. (= The speaker supposes that the referent is 'known' for the listener.) Therefore:

(21)\* Before I could talk to him, a policeman turned away from me.

The indefinite article indicates here that 'a policeman' is 'unknown' and for the first time under discussion. 'him' cannot have as its 'antecedent' somebody who is introduced on the scene first after the pronoun has been used.

Derek Bickerton treatment of 'pronominalization' (Bickerton 1975) is radically different from all the above-mentioned works. He has completely rejected the old syntactic rules of pronominalization and tries to use notions similar to those of FSP: "Pronominalization flows bidirectionally, and across sentence boundaries, from presupposed to asserted NP, and between presupposed NP, except where one NP has been presupposed throughout its derivational history and the other has not, in the latter case, pronominalization shall be from the more-consistently to the less-consistently presupposed." (Bickerton 1975, p. 32-33). However, his instrument, i.e., the sentence bipartition into 'presupposition' and 'assertion' (in FSP terms: the context dependent part and the context independent part) is insufficient.

As Bickerton's hypothesis is based on various constructions which were extremely difficult to explain within the traditional framework (above all the constructions where rhematicity is expressed syntactically - for example the types:

It was + rheme + who + theme (It was PETER who killed his wife.)  
 What + theme + was + rheme (What annoyed Bill was my punching him.),  
 he can accommodate these constructions which consist of sharply delimited context-dependent and context-independent parts. However, there are many sentences where such delimitation is not possible, e.g., sentences which consist of context-independent elements only. Bickerton's rule cannot account for the following examples either:

(What happened with John?)

(22) \*He was killed when John tried to escape.

(23) The cops killed him when John tried to escape.

It is impossible to see any difference in 'presuppositional consistency' between 'he' in the first and 'him' in the second sentence.

Bickerton seems to have completely missed the fact that there do exist sentence-stressed 'asserted antecedents' (in FSP terminology: antecedents which are the most rhematic elements). For example, our sentence (12) (It was PETER who killed his wife.). There are some dubious cases in his reasoning, too. It is not clear why, e.g., 'Rhoda' in 'Rhoda's apartment' is claimed to be asserted and 'apartment' is presupposed - Bickerton claims that 'Rhoda's apartment' is derived from 'The apartment belongs to Rhoda.', while the more usual interpretation 'Rhoda has an apartment' where 'Rhoda' is 'presupposed' and 'apartment' is 'asserted' seems as plausible as Bickerton's claim.

#### Coreference Rules Stated in Terms of FSP<sup>3</sup>

A) All pronouns must have antecedents in the preceding linguistic context or the situation.<sup>8</sup>

B) Coreference ('pronominalization'), both 'forward' and 'backward', is possible only when the degree of Communicative Dynamism (CD)

carried by the 'antecedent' is not substantially higher than that carried by the pronoun(s). This rule is subject to the following restrictions:

- C) The 'antecedent' can be a rheme proper in sentences of second-in-stance level (see Appendix A) if coreference is possible in the first-instance level sentence from which the second-in-stance level sentence is derived.
- D) No pronominal rheme proper is allowed before the 'antecedent'.<sup>9</sup>
- E) The 'antecedent' is obligatory as the theme proper.<sup>4</sup>

Let's exemplify these rules:

Rule A

A has been already demonstrated in (21), which is repeated here for convenience:

(21) \*Before I could talk to him, a policeman turned away from me.

Some more examples:

(24) He is stupid.

(25) (Who do you suppose I calmed when I saw him getting mad?) \*I calmed him before Harry did something rash.

(24) would be without meaning if the speaker did not refer to a male person. It is always the speaker/writer who determines whether there is an antecedent, but if the listener/reader cannot find an antecedent in the context/situation, he reacts: (A: He is stupid. B: Who?<sup>5</sup>).

As (25) is preceded by a rhetorical question, the speaker cannot presume that the listener knows the antecedent. As for the following dialogue, it is possible only when the speaker B presumes that A closely connects John with Mary, so that A immediately identifies the referent of the pronoun.

A: Tell me about John. B: (26) Since I despise her, I have not seen

John and Mary for ages.

Rule B

(3) John left town after he robbed the bank.

(4) \* He left town after John robbed the bank.

(5) After John robbed the bank, he left town.

(6) After he robbed the bank, John left town.

(3), (5), and (6) allow coreference as the antecedents/postcedent carry a low degree of CD. It is the linearity of (3) and (5) that determines the low degree of CD. As for (6), even there 'John' belongs to the theme (because of Rule A - 'John' must be context dependent and therefore thematic. If 'John' were context independent, then it could not be coreferent with 'he' which demands the existence of an antecedent). (4) does not allow coreference. The pronominal subject of (4) is definitely the theme proper, while 'John' in the subordinate temporal clause placed after the main clause (= a rhematic subclause) belongs to the rheme. It is not enough that 'John' carries a low degree of CD within the rhematic part. The difference in degrees of CD is too great, so coreference is impossible.

(27) \* It disturbs her that Mary is pregnant.

(28) It disturbs Mary that she is pregnant.

(29) That Mary is pregnant disturbs her.

(30) That she is pregnant disturbs Mary.

As has been shown in Svoboda (1968), there are two possible FSP interpretations of sentences with extraposition. Either the subclause is thematic and VP of the main clause is rhematic, or they are both rhematic and the subclause is the rheme proper. In (27) the former possibility is out of question because the unstressed 'her' cannot belong to the rheme, and even the latter interpretation is modified: The subclause is rhematic and 'her' is thematic for the same reason as in the

first case. The thematic pronoun (the theme proper in this case) cannot be coreferent with 'Mary' which belongs to the rheme. As for (28), the second one of the above-mentioned interpretations allows coreference. In (29), pronounced in the 'normal', 'unmarked' way (with sentence stress on 'disturbs'), both the subclause (and therefore even 'Mary') and the pronoun are thematic; coreference is possible. In (30) either 'Mary' is unstressed (and bears a low degree of CD) - then coreference is possible - or it carries sentence stress, then the sentence in question comes from a sort of 'presupposition' (a first-instance level sentence) like 'That she is pregnant disturbs somebody.' and 'Mary' identifies the 'somebody' - coreference is possible, too (Rule C and B). (Of course, the first-instance level sentence must have an antecedent for 'she' (Rule A), and therefore it demands a context where the antecedent does exist, for example, we have to know that a certain girl/some girls of a limited number is/are disturbed by her/their pregnancy.)

#### Rule C

Naturally, all the previous sentences may get a lot of various contrastive stresses. Such sentences belong from FSP's point of view to the second-instance level. The heavily stressed word/phrase is rhematic and the rest of the sentence is thematic. But even then it is possible to speak about different degrees of CD within the first-instance level sentence from which the second-instance level sentence originates. (And the same is true even for sentences in real contexts.)

(31) John DID leave town after he robbed the bank.

(32) John left town after he robbed the BANK.

(33) JOHN left town after he robbed the bank.



(34) It was JOHN who left town after he robbed the bank.

(35) \*HE left town after John robbed the bank.

(36) \*He left town AFTER John robbed the bank, etc.

Coreference is possible in (31), (32), (33), and (34) (Rule B and Rule C - for example, (33) and (34) come from a first-instance level sentence which is something like 'Somebody left town after he robbed the bank.', where coreference is possible (Rule B), and 'JOHN' identifies 'Somebody'.) In (35), it is not possible to get coreference reading in the first-instance level sentence ('Somebody left town after John robbed the bank.), therefore coreference is not possible in the second-instance level sentence, either. (Besides, (35) would have clashed with Rule D, too.) (36) does not allow coreference as there is a great difference in CD between the pronoun and 'John' even within the first-instance level sentence (Rule B), therefore coreference is impossible even within the second-instance level sentence (Rule C). The same rules (Rule B and Rule C) determine whether coreference is possible or not in (37) - (40):

(37) \*It disturbs HER that Mary is pregnant. (Cf. (35))

(38) \*It disturbs her that Mary IS pregnant. (Cf. (36))

(39) It disturbs MARY that she is pregnant. (Cf. (33))

(40) That Mary is pregnant disturbs HER (only).

Coreference is possible in (40). (Rule B)

As it has been mentioned, 'backward pronominalization into main clause' is possible at least for some speakers when 'the resulting pronoun' is not the subject of the main clause (Lakoff 1968). But this is true in certain cases only, e.g. in (23). The difference in the degree of CD between the pronoun and the NP which is supposed to be coreferent with it, is not so great and that's why coreference is possible, at least

for those speakers who are more sensitive to the various 'shades' of CD. (According to Firbas' theory, an unstressed pronominal object is thematic but it carries higher degree of CD than a pronominal subject.) On the other hand, coreference is impossible in (41). 'John' may be stressed or unstressed but it belongs to the part of (41) which carries the highest degree of CD. Coreference is ruled out by Rule B.

(23) The cops killed him when John tried to escape.

(41) \*The cops killed him when they tried to arrest John.

Naturally, coreference is out of question in (22) since 'He' is the theme proper and the degree of CD carried by 'John' is in relation to it too high (Rule B).

(22) \*He was killed when John tried to escape.

Even (42) - (47) confirm our Rule B:

(42) They are not very clever, these politicians.<sup>6</sup>

(43) He would have been like a son to both of us if we could have kept John away from the influence of his friends.<sup>6</sup>

(44) John said that he was safe.

(45) \*He said that John was safe.

(46) He was safe, John said / said John.

(47) \*John was safe, he said / said he.

'these politicians' of (42) is a mere repetition of the thematic 'they' regardless of its final position. Of course, one could keep the usual 'precede and command' rule if one claimed that the surface structure of (42) is such that 'they' does not command 'these politicians', but nothing similar can help in (43) and (47). (43) cannot be explained with any other proposed rules, but it is okay according to our Rule B. 'John' in (43) is definitely not too rhematic - the subclause can be interpreted as a sort of parenthetical clause which follows the highly

rhetic 'a son to both of us' and even in the subordinate clause some elements carry higher degree of CD than 'John'. The 'strange' (47) does not allow coreference as 'he' is the theme proper both in (54) and (47). The distribution of CD which is determined by the semantics of these sentences, is not changed by the different word orders. 'he said / said he' in (47) is a sort of 'parenthetical theme' and the difference in the degree of CD between 'he' and the rhematic 'John' is too great.

Some more examples:

(48) \*He believes that Hubert will win big in '72.

(49) \*Mary told him that Hubert would take it all in '72.

(50) \*We gave him a picture of Hubert.

(51) His mother believes that Hubert will win big in '72.

(52) Mary told his mother that Hubert would take it all in '72.

(53) We gave his mother a picture of Hubert.

In (48) - (50) coreference is impossible as 'he' resp. 'him' carries too low a degree of CD. A little increase of CD carried by the pronoun (sentences (51) - (53)) makes coreference possible. For example, the pronominal subject of (48) is the theme proper; besides, it is (being a pronoun) the theme proper with the lowest possible degree of CD. 'His' of (51) belongs to the NP which carries the lowest degree of CD in the sentence, however, 'His' carries higher degree of CD than 'He' of (48).

As (53) shows, the 'antecedent' can carry a relatively high degree of CD if the pronoun carries a high degree of CD, too. (Our Rule B again.) In (50) coreference is impossible as the difference in CD between the pronominal object (the element with the next lowest degree of CD in (50)) and the 'antecedent' which belongs to the rheme is too great. The

pronoun of (53) carries a little higher degree of CD and coreference becomes possible.

Our rules are valid not only for whole sentences but also for their parts:

(54) \*Mary's telling him that Hubert would win was a major mistake.

(55) Mary's telling his mother that Hubert would win was a major mistake.

In relation to the whole sentence, 'Mary's telling him that Hubert would win' is thematic but when coreference possibility is judged within the cycle in question, the result is the same as in (49). Naturally, coreference is possible in (55) as it has been in (52).

The rules which has been exemplified so far are A), B), and C). What about D) and E)? Rule D prohibits coreference in sentences like (56) and (57) (and (35), (37)):

(56) \*After HE robbed the bank, John left town.

(57) \*The cops killed HIM when John tried to escape.

Rule E is, in fact, just a special case of Rule B. All the previous sentences that begin with pronominal subjects of the main clause exemplify not only Rule B but also Rule E. The pronominal subject of the main clause placed as the first word of the sentence cannot be anything else than the theme proper, but this is prohibited by our Rule E, which says that the difference in the degree of CD between a pronoun which is the theme proper and its 'antecedent' is too great.<sup>6</sup> That's why only (58) is acceptable in the following dialogue and (59) is not:

A: What will John do tomorrow?

B: (58) If he can, John will go to see a movie.

(59) \*If John can, he will go to see a movie.

In another context is (59) okay:

A: Will John go to see a movie if he can?

B: (58b) If he can, John will certainly go.

(59b) If John can, he will certainly go.

(The semantics has been slightly changed to get more natural answers.)

Both (58b) and (59b) are okay since 'if X can' and 'X' are equally context dependent (and therefore thematic). In such case we can choose any of the Xs as the 'antecedent'.

As for (43), it is not a counterexample to Rule E. We are talking about 'us' (about our relations to John) and about 'John'. 'John' is not the whole theme proper - it is just a part of it. That's why it can be 'pronominalized' in accordance with Rule B.

Rule E can also explain the 'strange' sentences with topicalized adverbials:

(60) In her apartment, Mary smokes pot.

(61) \*In Mary's apartment, she smokes pot.

(62) In Mary's apartment, her brother smokes pot.

(63) ?? In Mary's brand new apartment in Detroit, she smokes pot.

(64) Near her, Mary found a snake.

(65) \*Near Mary, she found a snake

(60) is okay since 'Mary' is the theme proper. (61) does not allow co-reference as it is impossible to get some other theme proper than 'she'. Neither 'Mary' which belongs to the adverbial that carries higher CD than 'she' nor 'apartment' can be the theme proper. As for (62), any of the two genitives suits as the theme proper. (Even 'In her apartment, Mary's brother smokes pot.' is okay.) (63) is acceptable for some people because the structure of the long adverbial phrase strongly resembles that of a clause, where 'Mary' can become

the theme proper of the whole sentence. The explanation of (60) and (61) is valid even for (64) and (65). The adverbial carries higher degree of CD than the subject, which is the theme proper. (It is obvious that 'Near' does not suit as the theme proper, either.)

#### Some 'Bad Cases of Pronominalization'

The above-mentioned 'direct discourse analysis' of (7) (Kuno 1972a) was refuted by Hinds (1975a). The invalidity of Kuno's claim was proved by (8), (9), and (10). According to Hinds, 'pronominalization' is impossible when passivation had been used to indicate that a NP is 'rhematic'. Quotation marks are used because, as it has been mentioned already, Hinds' concept of theme and rheme has only a certain superficial similarity with the terminology of Prague school.) However, Hinds is wrong even within his own theory, as (66) shows. What is wrong with (7) when (8), (9), (10), (66), (67), and even (68) are okay?

- (7) \*That John was the best boxer in the world was claimed by him.  
 (8) That John was the best boxer in the world was never claimed by him.  
 (9) That John was the best boxer in the world was loudly and repeatedly claimed by him.  
 (10) That John was the best boxer in the world was claimed by him but nobody would believe such nonsense.  
 (66) The woman who rejected Peter is hated by him.  
 (67) That John was the best boxer of the world was claimed by HIM.  
 (68) That John was the best boxer of the world was denied by him.

(7) is bad because the theme-rheme structure it indicates does not make sense. The subordinate clause is thematic and the unstressed pronoun in the by-phrase is thematic, too. The only possible rhematic element is the remaining participle and copula. But it seems difficult (without contrastive stresses, i.e. without reading the sentence in question as

belonging to the second-instance level - for example: That John was the best boxer in the world WAS claimed by him.) to find a sensible context where the copula or the participle could be the rheme proper. This has nothing to do with coreference. (69) sounds as bad as (7). (69) ?? That John was the best boxer in the world was claimed by her (with 'her' unstressed).

(8), (9), (10), (66), (67), and (68) are okay since the VPs are made naturally rhematic with the rhematic adverb 'never', with the modification of the verb with 'loudly and repeatedly' etc. (67) has the rhematic 'HIM'. (66) and (68) contain participles which can be rhematic in a natural way - e.g. 'deny' implies negation and negation is usually the most rhematic part of a sentence.

Our rules can also explain sentences (70) and (71). It is usually said that coreference is possible in (70) and impossible in (71). It is, in fact, true only for the most usual ways of reading of these two sentences, i.e. (20) resp. (77). (70) can be pronounced, for example, as (20), (72), (73). (80) can be pronounced as (74) - (77).

(70) His brother visited John.

(71) His brother was visited by John.

(20) His BROTHER visited John.

(72) \*HIS brother visited John.

(73) His brother visited JOHN.

(74) His BROTHER was visited by John.

(75) \*HIS brother was visited by John.

(76) His brother WAS visited by John.

(77) His brother was visited by JOHN.

Everything is as our rules predict. The 'antecedent' carrying a relatively low degree of CD makes coreference possible in (20), (74), (76),

and coreference is impossible in (72) and (75), where our rule D prohibits it. As for (73) and (77), which are not acceptable 'normally', they become acceptable in a context in which the issue is whether John visited his brother or his brother him. In such a context (i.e. as sentences of second instance level), (73) and (77) allow coreference (Rule C). Finally there are two types of sentences for which Postal claims the existence of his Global Constraint on Pronominalization and the restriction prohibiting 'backward pronominalization' across a copular verb of referential identity.

(78) Who hates his wife?

(79) \*Who does his wife hate?

(80) Who is hated by his wife?

(81) It was Peter who killed his wife.

(82) \*It was Peter who his wife killed.

(83) It was Peter who was killed by his wife.

(84) His wife hates who?

(16) It was my punching him that annoyed Bill.

(85) It was my punching Bill that annoyed him.

(15) \*What annoyed him was my punching Bill.

(86) What annoyed Bill was my punching him.

(78), (80), (81), and (83) allow coreference. These sentences are second-instance level sentences coming from 'presuppositions' like 'Somebody hates his wife', 'Somebody is hated by his wife' etc. (Rule C and Rule B). In (79) and (82), coreference is impossible since the first instance level sentences, which are of following types: 'His wife hates somebody' and 'His wife killed somebody', do not allow coreference. These 'presuppositions' are wrong as the given 'wife' is falsely 'specified' as belonging to an unspecified person (Rule A).



(87) does not allow coreference, either. The first-instance level sentence would be 'He was bitten by somebody's dog,' which cannot contain any intrasentential coreference. 'He' must have some earlier antecedent (Rule A) and the known antecedent cannot be coreferent with the unspecified 'somebody'.

(87) \*He was bitten by PETER's dog.

(One should be conscious of the difference between the semantic rules of coreference and the possibilities given by the semantics of a given sentence and our knowledge of the world. For example, (77) is okay, while (88), which has the same theme-rheme structure (as well as deep structure and surface structure) does not allow coreference since it is practically impossible to imagine a context (a first-instance level sentence) where the question is who was eaten by whom. If we succeed in imagining a suitable context, coreference becomes possible.

(88) ? Their keepers were eaten by the TIGERS.)

As for the possible reading of (78) - (80) and even (84) as incredulity questions, the coreference interpretation is possible, probably for the majority of speakers, because such questions do not correspond with their distribution of CD to the statements they are 'echoes' to. 'Who' in the incredulity questions does not demand an answer, such a 'who' stands for an already known NP and the speaker just shows that he is surprised that the statement which has been made is valid for the NP in question. Naturally, there may be speakers who evaluate the theme-rheme structure of such sentences according to some more 'formal' criteria ('who' in all questions is rhematic for them), and for them (79) is wrong as an incredulity question, too, not to mention (84) which is 'completely wrong', the obligatory placing of 'who' at the beginning of the sentence and Do-Insertion not having been accomplished. (In fact,

there is another sentence type which is wrong for some speakers: 'This is the man who the fact that he had cancer surprised.' The possibility or impossibility of coreference depends again on the speaker's evaluation of the degree of CD carried by the relative 'who' - Rule B again.) In (16), the syntactic construction guarantees that 'Bill' is thematic and coreference is possible (Rule B), except for (16) with sentence stress on 'him' (Rule D). In (85), coreference is possible if sentence stress is placed on 'my' or 'punching', (Rule B), but if it is placed on 'Bill', coreference is impossible (Rule C). (15) prohibits coreference as our Rule B predicts. (86) allows coreference as 'Bill' belongs to the theme (Rule B).

### Conclusion

The rules that have been presented and exemplified in this paper can do without the whole heterogenous collection of rules one was forced to accept otherwise. The new rules are simpler, fewer and homogenous.<sup>7</sup> They are also interesting because they show that the theory of FSP, which is ignored by the majority of linguists (or - if accepted - which is 'man-handled' in unbelievable ways), is an instrument necessary for explanation of various linguistic phenomena. Even the validity of the analytical procedures used in FSP is indirectly confirmed by their use which enables us to make correct predictions concerning coreference relations.

Just to name some other cases where FSP can explain certain phenomena which cannot be understood otherwise, or which are usually 'explained' by placing a 'name label' on them: Fillmore (1970) gives the following sentences and cannot do more than note that the last one is ungrammatical.

(89) An oak developed out of every acorn.

(90) Every acorn developed into an oak.

(91) Every oak developed out of an acorn.

(92) \*An acorn developed into every oak.

The indefinite article of a thematic subject means either 'one of the', or 'a single', or it can signal a generic noun. None of these meanings fits in (92); it is impossible (because of the semantics of (92)) to interpret 'An acorn' as generic, or 'one of the acorns', or 'a single acorn'. It is quite unnecessary to try to speak about different 'scopes' of the quantifiers.

A similar problem was brought up by Lakoff (1970)

(93) Johan and Max saw an explosion.

(94) An explosion was seen by John and Max.

(93) means either 'John and Max saw an explosion and it was the same explosion', or 'John saw an explosion and Max saw another one.'. (94) is said to mean only 'John and Max saw an explosion and it was the same one.' The difference in possible meanings of (93) and (94) is usually explained with different 'scopes' of the indefinite articles in the active and passive sentence. But this is not true. If (94) is pronounced with sentence stress on 'explosion', it can be interpreted in both ways. Then, those who believe in 'scopes' would have to formulate an additional rule for certain sentence stresses which can change 'scopes'. The indefinite article of the thematic subject of (94) cannot obviously mean that the 'explosion' was generic and the other two meanings ('a single' and 'one of the') limit the number of explosions to one. However, if 'explosion' of (94) carries sentence stress, it cannot be thematic and the indefinite article with a rhematic noun means 'some', 'an unspecified'. That's why both interpre-

tations are then possible. The same is true even for (93). The rhematic NP with the indefinite article allows both interpretations. And vice versa - (95) can mean only 'John and Max saw an explosion and it was the same one.' as the thematic 'explosion' with the indefinite article must mean 'a single explosion' or 'one of the explosions'.

(95) JOHN and MAX saw an explosion.

To sum up: there may be reasons to formulate rules of the semantic interpretation of the sentences above in terms of 'scopes', but only if these are based on theme-rheme structure, not on sentence linearity only (as the generally accepted mechanical 'scopes' are).

#### Appendix A

"The concept of communicative dynamism is based on the fact that linguistic communication is not a static, but a dynamic phenomenon. By CD I understand a property of communication, displayed in the course of the development of the information to be conveyed and consisting in advancing this development. By the degree or amount of CD carried by a linguistic element, I understand the relative extent to which the element contributes to the development of the communication, to which, as it were, it 'pushes the communication forward'. Thus, if examined in its unmarked use, the sentence He was cross could be interpreted in regard to the degrees of CD as follows. The lowest degree of CD is carried by He, the highest by cross, the degree carried by was ranking between them."

(Firbas 1971, p. 135-136)

"We hold that between the comparatively least important element, the theme proper, and the comparatively most important element, the rheme proper, one can observe a gamut of degrees of varying importance, of varying communicative value, of varying CD. It is, of course, not always easy to draw an exact dividing line between the transition and the theme on the one hand, and the transition and the rheme on the

other. In such cases it is necessary to attempt at least a correct estimate of the relative importance of the elements composing the analysed structure."

(Firbas 1959, p. 42)

"In this connection it is worth notice that in some cases even a thematic element may contribute considerably to the development of discourse. Thus in a sentence of the type A girl broke a vase, the thematic subject carries a comparatively high degree of CD (the non-generic indefinite article marking out a new idea). Yet as the other elements are more dynamic still, the subject is felt to be thematic owing to the pressure exerted by the basic distribution of CD. This means that the fact of the theme carrying the least amount of CD does not preclude the possibility of its carrying a new piece of information."

(Firbas 1959, p. 42-43)

"I believe that much valuable light can be thrown on the function of language in the very act of communication by a consistent inquiry into the laws determining the DISTRIBUTION of degrees of CD over linguistic elements capable of carrying them.

The following note will be relevant here. It will answer the question of what linguistic elements can become bearers of degrees of CD. I agree with D.S. Worth that a linguistic element - sentence, noun phrase, word, morpheme, submorphemic segment, etc. - may be singled out in order to establish a sharp ad hoc opposition (contrast): John WAS winning. Jenda VÍTĚZIL. The fact that VÍTĚZIL may appear in sharp ad hoc contrast to vítĚzí ('is winning') shows that the element -il may become the actual bearer of the contrast. Under the circumstances it would be the only element conveying new information (and therefore be contextually independent), whereas all the other elements would convey known information (and therefore be contextually dependent). Normally, the element -il would not carry the highest degree of CD as it does in the highly marked example under discussion; but even then, on account of its semantic content, it would have to be regarded as a carrier of CD. All linguistic elements, including morphemes and submorphemic exponents, are capable of carrying degrees of CD, as long as they convey some meaning. It may have been gathered from the above note that elements become con-

textually dependent and in consequence carriers of the lowest degrees of CD owing to the operation of the context. They assume this function irrespective of the positions they occupy within the linear arrangement. (I avoid the term 'word order' here, because words are not the only elements concerned.) Strictly speaking, contextual dependence or independence is determined by what I have called the narrow scene, i.e. in fact the very purpose of the communication. Thus in the sentence John has gone up to the window, the window may be well known from the preceding context, but the purpose of the communication being the expression of the direction of the movement, the window necessarily appears contextually independent. Under the circumstances, it is - to use Halliday's appropriate terms - non-derivable, non-recoverable from the preceding context.

Let me now turn to contextually independent elements. In determining their degrees of CD, two further factors are in play: (i) the semantic structure, (ii) positions of the elements within the linear arrangement. By the semantic structure of a sentence I understand the semantic contents of the sentence elements and the semantic relations into which they enter.

I will first illustrate the operation of the semantic structure. An object expresses the goal (outcome) of an action conveyed by the accompanying verb. Provided it is contextually independent, it will carry a higher degree of CD than the verb. This is because from the point of view of communication, an unknown goal (outcome) of an action appears to be more important than the action itself. A contextually independent object will carry a higher degree of CD than the verb irrespective of the positions occupied within the linear arrangement. (I have read a fine book, Ich habe ein schönes Buch gelesen, Četl jsem pěknou knihu.)

Similarly, a contextually independent adverbial element of place expressing the direction or destination of a motion will exceed in CD a verb expressing the motion. This is because, communicatively speaking, an unknown direction or destination of a motion is more important than the motion itself. (I flew to London, Ich flog nach London, Ich bin nach London geflogen, Letěl jsem do Londýna.)

As to the subjects of the two structures discussed in the two preceding paragraphs, each will carry the lowest degree of CD within the sen-

tence provided at least one of the remaining two elements is contextually independent. This is understandable, for a known or unknown agent expressed by the subject appears to be communicatively less important than an unknown action expressed by the verb and/or an unknown goal (expressed by the object or the adverbial element of place) at or towards which the action is directed. (Cf. the examples offered above and also A girl was reading an interesting book, A girl was travelling to an unknown town, Ein Mädchen hat ein interessantes Buch gelesen, Ein Mädchen ist nach einer unbekannten Stadt gereist.)

The situation would be different if the subject expressed a person or thing 'existing' or 'appearing on the scene' and if it were accompanied by a verb expressing the notion of 'appearance' or 'existence on the scene'. If under these circumstances it is contextually independent, the subject will carry a higher degree of CD than the verb. This is because, communicatively speaking, an unknown person or thing appearing on the scene is found to be more important than the fact of existence or act of appearing itself. This holds good irrespective of the positions occupied by the respective elements within the sentence. (A girl came into the room, Ein Mädchen kam ins Zimmer, Ins Zimmer kam ein Mädchen, Ins Zimmer ist ein Mädchen gekommen, Do pokoje vešla dívka.)

Not all semantic contents and relations, however, are capable of signalling degrees of CD in the way indicated above. There are evidently also such types of semantic content as let the linear arrangement itself determine the degrees of CD. Thus a contextually independent infinitive of purpose carries a lower degree of CD when occurring initially than when occurring finally (In order to see him, he went to Prague, He went to Prague in order to see him). Similarly, with the indirect and direct object, provided they are contextually independent, the one coming later within the linear arrangement carries a higher degree of CD. (He gave a boy an apple, He gave an apple to a boy.)

The notes that have just been offered have shown that the distribution of degrees of CD over the sentence elements (the signalling of the degrees) is an outcome of an interplay of three factors: context, semantic structure, linear arrangement. The notes have also indicated that according to the contextual situation, in other words, the contextual dependence, the distribution of degrees of CD over a sentence struc-

ture may vary. All the possible variations (realizations of the distribution) constitute the contextual applicability of the sentence structure. Before proceeding further, let me insert a note on the linear arrangement.

Sentence linearity is an indisputable fact. It makes the speaker/writer arrange the linguistic elements in a linear sequence, in a line, and develop the discourse step by step. I believe to be right in assuming that the most natural way of such gradual development is to begin at the beginning and proceed in steady progression, by degrees, towards the fulfilment of the communicative purpose of the discourse. If this assumption is correct, then a sequence showing a gradual rise in degrees of CD (i.e. starting with the lowest degree and gradually passing on to the highest degree) can be regarded as displaying the basic distribution of CD. I also believe to be right in assuming that this conclusion is quite in harmony with the character of human apprehension.

On the other hand, it seems to be equally in accordance with the character of human apprehension that in a discourse made up of a longer string of verbal sentences, the basic distribution of CD in the fullest sense of the word (i.e. one throughout which, gradually, every element becomes a carrier of a higher degree of CD than its predecessor) can practically never be accomplished. Within such a distribution every element would convey new information. But this is not the way the discourse is structured. In order not to jeopardize comprehension, the discourse is continually interspersed with elements conveying information known (derivable, recoverable) from the previous context, i.e. by elements that have been designated here as contextually dependent. It is through these elements that relief from the steady flow of new information is constantly provided.

The field within which the distribution of CD takes place is naturally the entire discourse. This field, which may often become very extensive, is subdivided into fields of lower rank, provided, e.g., by chapters, paragraphs, sentences, subordinate clauses. Remaining within the sphere of complex sentences and structures ranking below them, I subscribe to A. Svoboda's view that distributional fields are provided by grammatical structures that convey either explicit (open) or implicit (hidden) predication. (Under the latter heading come structures formed by headwords and their accompanying attributive words or phrases.) In provid-



ing distributional (communicative) fields, grammatical structure cuts, as it were, longer or shorter sections out of the linear flow of the discourse. As may have been gathered from what has been said before, according to their contextual dependence distributional fields may function in different perspectives. As the sentence is not the only type of structure providing a distributional field, it is - strictly speaking - possible to speak of other kinds of functional perspective than that of a sentence, e.g., that of a subordinate clause or that of an attributive construction.

The mutual relations between context, semantic structure and linear arrangement may be summed up as follows. Determining the contextual dependence of the distributional field, the context overrules the semantic structure and the linear arrangement in the interplay of means signalling the degrees of CD. It will be remembered that the semantic structure and the linear arrangement (in other words, the basic distribution of CD) can effectively perform the signalling function only within that section of the distributional field which has remained unaffected by the context, i.e. that section which has remained contextually independent."

(Firbas 1971, p. 136-139)

"Semantic structure ceases to operate in what we have termed after D.L. Bolinger second instance sentences, i.e. in such as contain one element singled out for special attention (usually for the sake of heavy contrast) and functioning as one-element rheme proper, all other elements forming an extensive theme proper. Any element can become rheme proper within second instance (HE wrote an interesting book)."

(Firbas 1966, p. 241)

" . . . we maintain that the function of the sentence in the act of communication can be successfully interpreted if three levels are kept separate: those of the semantic and the grammatical structure of the sentence and that of FSP. As we see it, the theory of FSP makes it possible to understand how the semantic and the grammatical structure of the sentence function at the very act of communication, i.e. at the moment they are called upon to convey some extra-linguistic reality reflected by thought and are to appear in an adequate kind of perspective."

(Firbas 1966, p. 241)

It is important to stress the great difference between the Prague school's FSP and the use of the terms 'theme' and 'rheme' by Halliday (1967-68) or Enkvist (1974), whose 'themes' and 'rhemes' are synonymous with initial resp. final sentence elements. The Prague concept of CD is a generalization and an abstraction from analyses of concrete texts, where it is possible to divide every sentence into the known ('given', context dependent) part and the unknown ('new', context independent) part, a generalization and an abstraction which is valid even for the initial sentence of a discourse, where all information may be 'new' ('Once upon a time there was a king. '), and isolated sentences without any known context.

#### Appendix B (Two Remarks on Reflexives)

It is rather difficult to discuss the earlier attempts to explain English reflexives since nearly all crucial examples in Jackendoff (1972) or Postal (1971) are semantically dubious - e.g. 'I sold the slave myself.' or 'I bought the slave for himself.' Let's take just one reasonable 'sample':

\* John was shaved by himself. (with unstressed 'himself')

Postal (1971) proposes for sentences like this a constraint called the Cross-Over Principle, which says roughly that a transformation cannot move an NP over another NP with which it is coreferent. However, this cannot explain why the same sentence is okay with stressed 'himself' and Postal has to claim that the above-mentioned sentence with stressed 'himself' has the following 'fancy' Deep Structure:

[the one [one shaved John]<sub>S<sub>1</sub></sub> was John]<sub>S<sub>2</sub></sub>

Of course, it is impossible to prove that a theory which uses such powerful instruments as global constraints and which can postulate a new

abstract deep structure every time it meets a problem, is wrong, but it would be nice to do without such things.

Jackendoff's solution is based on his so-called Thematic Hierarchy Condition on Reflexives (this has nothing to do with the notion of 'theme' of FSP) which operates with semantic roles vaguely reminiscent of Fillmore's 'cases'. As another version of Jackendoff's Thematic Hierarchy Condition (on the passive transformation) is wrong (as Gee (1974) has shown) it would be rather optimistic to believe that this one does work, and Jackendoff does not make any attempt to explain why there should be any difference between the effect of his condition on the sentence with an unstressed and a stressed 'himself', either.

I believe that even the behaviour of reflexives can be explained within FSP. Our sentence with unstressed 'himself' is ungrammatical since it makes no sense from the point of view of FSP: The deep structure 'Himself shaved John' (which is, of course, wrong and which can be 'saved' by a passive transformation 'triggered' by the needs of FSP) cannot become the above-mentioned sentence with unstressed (and therefore thematic) 'himself' which would annul the result of the passive transformation. On the other hand, the stressed (and therefore rhematic) 'himself' does not clash with the needs of FSP, which triggered the passivization, and the sentence with stressed 'himself' is okay.

Harada - Saito 1971 has shown that there exist sentences which cannot get the right reflexive interpretation in the cyclical way Jackendoff 1972 proposes. In 'John believes himself to be hard for Bill to understand.', 'himself' must get coreferent with 'Bill' in the lower S it comes from and the correct coreferent interpretation (John + coref. himself) is blocked. But it is the same thing with all NP's which should be interpreted as coreferent but cannot be in Jackendoff's way, as Jackendoff

reasons about the cases of 'interpretatively asserted coreference' only and ignores the cases of 'sententially asserted coreference' (e.g. The morning-star and Venus are the same celestial body.) and 'factual coreference' (coreference based on our knowledge of the world, which makes it possible to 'mark as coreferent' e.g. the following NP's: 'the infamous American president at the beginning of 1970's' and 'Richard M. Nixon').

The solution may be in running the interpretation in cycles parallel to the syntactic transformational cycles with the possibility of 'loops' - e.g., reflexives would be first interpreted on their respective cycles and then, if they are raised, on the next cycle. If the reading on a higher cycle clashes with that obtained on the lower cycle the latter is annulled. This solves the problem presented in Harada - Saito 1971. Some kind of 'reinterpretative rule' is necessary even for the cases of 'sententially asserted coreference' and 'factual coreference' where the NPs in question become first marked as noncoreferent by the rules of 'interpretatively asserted coreference'.

Notes

This paper is an enlarged and revised version of 'Pronominalization Rules' published in Papers from the Third Scandinavian Conference of Linguistics, ed. by F. Karlsson, Turku 1976.

- 1 Wasow (1975) argues against the proposal of Harman (1972) who answers in Harman (1976). However, Harman (1976) avoids the argument presented above and his reasoning is far from convincing. For example, one of Wasow's arguments is: It is well known that the operation of Dative Movement depends on whether the direct object is pronominal. (\*He gave me it. vs. He gave me the book.) So Dative Movement must follow NP Placement. On the other hand, Passive can follow Dative Movement; and, according to Harman's proposal, passivization must precede the substitution of variables (NP Placement) - cf. the example with 'burglars'. These assumptions are, of course, jointly inconsistent. To save his theory, Harman (1976) is forced to 'cheat' in the following way:

- a) x gave y to z
- b) x gave NP to z (= NP Placement of one of the variables only)
- c) x gave z NP (= Dative Movement)
- d) z was given NP by x (= Passive which moves only variables)
- e) NP was given NP by NP (= NP Placement of the remaining variables).

This is no honest solution and it does not work either, it can save the crucial examples with 'burglars' only if the idea of transformation cycle is given up and there is no sensible way to generate sentences like 'The woman he loved deceived John.' as the pivotal point of Harman's proposal is that the leftmost occurrence of a variable is substituted by a NP while all other occurrences of that variable get pronominalized. However, Harman's proposal was quite revolutionary, as it was the first attempt to get rid of the 'precede and command' rule. For example, NP Placement applies to underlying structure of f) to yield (3):

f) [x left town [after x robbed the bank] S1 ] S2, [x:John]

(3) John left town after he robbed the bank.

Adverbial Clause Movement applies to (3) to yield (6):

(6) After he robbed the bank, John left town.

Alternatively, Adverbial Clause Movement applies to f) yielding

g) [After x robbed the bank, x left town], [x:John]

and NP Placement follows:

(5) After John robbed the bank, he left town.

There is, however, no way to get (4):

54) \*He left town after John robbed the bank.

- 2 There is a way to avoid the problems with NPs containing quantifiers. One can claim that only a N is pronominalized by another N instead of a NP by another NP. Then 'his' from Dougherty's example comes about from 'son + genitive'. Unfortunately for the transformationalists, stronger arguments against their hypothesis appeared during the seventies. For example, Kayne (1971) describes the behaviour of clitic pronouns in French. The argument is as follows: Clitic movement operates only on pronouns; therefore, if there is a transformational rule of pronominalization, it must precede cliticization. However, there are cases where clitic movement should then be able to move a pronoun to the left of its antecedent, but the resulting sentence is ungrammatical. It means that Constraint II demands that pronominalization follows cliticization. These incompatible demands can be avoided only by assuming that pronouns are present underlyingly and get interpreted later. Those sentences where there is no possible interpretation are ungrammatical.
- 3 These rules are based on what can be called 'a common denominator' of the categories previously used to limit the possibilities of 'pronominalization', which include a specification of sentence linearity, a specification of stress level, a definition of subject and nonsubject, a definition of main clause and subordinate clause, a specification of the sentence type (complex sentences with indirect speech versus other complex sentences), etc., etc.
- 4 Rule E is just a special case of Rule B. It specifies one typical case where the degree of Communicative Dynamism carried by a full NP would be too much higher than that of a pronominal theme proper. The restriction as it is formulated in Rule E is, in fact, to severe. As we shall see later, sentences (42) and (43) show that it is possible to have pronominal themes proper if their respective antecedents are placed in some kind of parenthetical expression which carries a very low degree of CD.

- 5 There is a usual sort of dialogue which has been observed by everybody: Person A is sitting alone in an otherwise empty room. Person B pokes his head in and asks: "Is he in here?". A replies: "Who?" regardless of the logical truth that the answer should be an instant "No." for any possible antecedent of "he".
- 6 See Note 4.
- 7 The interpretative theory has been used as it is not as discredited as the more usual transformational hypothesis. However, it must be emphasized that the conclusions are in no way dependent upon this theory. It would be possible to construct, for example, a model where pronouns are generated with indices determining coreference relations and the resulting sentences would be grammatical only when the coreference markers would not clash with the FSP interpretation of the transformations the sentences in question had undergone and their semantics.
- 8 and 9 follow after the bibliography.

#### Bibliography

- Akmajian, A. (1973), "The Role of Focus in the Interpretation of Anaphoric Expressions", in Anderson, S.A. and Kiperski, P. (eds), A Festschrift for Morris Halle, p. 215-226
- Akmajian, A. and Jackendoff, R.S. (1970), "Coreferentiality and Stress", Linguistic Inquiry 1:1, 124-126
- Bach, E. (1970), "Problominalization", Linguistic Inquiry 1:1, 121-122
- Bickerton, D. (1975), "Some Assertions about Presuppositions about Pronominalization", in Papers from the Parasession on Functionalism, 24-35
- Bresnan, J. (1971), "A Note on the Notion 'Identity of Sense Anaphora'", Linguistic Inquiry 2:4, 589-597
- Chomsky, N. (1957), Syntactic Structures  
 (1965), Aspects of the Theory of Syntax  
 (1970), "Remarks on Nominalizations", in Jacobs and

- Chomsky, N.                    Rosenbaum (eds), Readings in Transformational Grammar, 184-221
- (1971), "Deep Structure, Surface Structure, and Semantic Interpretation", in Steinberg and Jakobovits (eds), Semantics, 183-216
- Cole, P. (1974), "Backward Pronominalization and Analogy", Linguistic Inquiry 5:3, 425-443
- Culicover, P.W. (1976), "A Constraint on Coreferentiality", Foundations of Language 14:1, 101-118
- Daneš, F. (1964), "A Three-Level Approach to Syntax", Travaux linguistiques de Prague 1, 225-240
- (1974a) ed., Papers on Functional Sentence Perspective
- (1974b), "Functional Sentence Perspective and the Organization of the Text", in Daneš 1974a, 106-128
- Dougherty, R.C. (1969), "An Interpretative Theory of Pronominal Reference", Foundations of Language 5:4, 448-519
- Dvořáková, E. (1964), "On the English and Czech Situational Adverbs in Functional Sentence Perspective", Brno Studies in English 4, 129-142
- Emonds, J. (1973), "Alternatives to Global Constraints", Glossa 7:1, 39-80
- Enkvist, N.E. (1974), "Några textlingvistiska grundfrågor", in Teleman - Hultman, Språket i bruk, p. 172-206
- Fillmore, C.S. (1970), "Subjects, Speakers, and Roles", Synthese 21, 251-274
- Firbas, J. (1959), "Thoughts on the Communicative Function of the Verb in English, German and Czech", Brno Studies in English 1, 39-68
- (1964a), "From Comparative Word-Order Studies", Brno Studies in English 4, 111-128
- (1964b), "On Defining the Theme in Functional Sentence Analysis", Travaux linguistiques de Prague 1, 267-280
- (1966), "Non-Thematic Subjects in Contemporary English", Travaux linguistiques de Prague 2, 239-256



- Firbas, J. (1971), "On the Concept of Communicative Dynamism in the Theory of Functional Sentence Perspective", Sborník prací Filozofické fakulty Brněnské university A19, 135-144
- (1974), "Some Aspects of the Czechoslovak Approach to Problems of Functional Sentence Perspective", in Daneš 1974, 11-37
- Gee, J.P. (1974), "Jackendoff's Thematic Hierarchy Condition and the Passive Transformation", Linguistic Inquiry 5:2, 304-308
- Grinder, J. and Postal, P.M. (1971), "Missing Antecedents", Linguistic Inquiry 2:3, 269-312
- Halliday, M.A.K. (1967-68), "Notes on Transitivity and Theme in English", Journal of Linguistics 3, p. 37-81, 199-244; 4, p. 179-215
- Harada, S.I. and Saito, S. (1972), "A Non-Source for Reflexives", Linguistic Inquiry 2:4,
- Harman, G. (1972), "Noun Phrases Derived from Variable Binding Operators". To appear in Forms of Representation, ed. by Freed, Marras and Maynard
- (1976), "Anaphoric Pronouns as Bound Variables: Syntax or Semantics?", Language 52:1, 78-81
- Hinds, J. (1975a), "Passives, Pronouns, and Themes and Rhemes", Glossa 9:1, 79-106
- (1975b), "Backward Pronominalization across Coordinate Structures", Linguistic Inquiry 6:2, 330-335
- Jackendoff, R.S. (1968), "An Interpretative Theory of Pronouns and Reflexives". (Mimeographed by Indiana University Linguistics Club, Bloomington)
- (1972), Semantic Interpretation in Generative Grammar
- Karttunen, I. (1971), "Definite Descriptions with Crossing Coreference. A study of the Bach-Peters Paradox". Foundations of Language 7:2, 157-182

- Kuno, S. (1972a), "Pronominalization, Reflexivization and Direct Discourse", Linguistic Inquiry 3:2, 161-195
- (1972b), "Functional Sentence Perspective", Linguistic Inquiry 3:3, 269-320
- (1975), "Three Perspectives in the Functional Approach to Syntax", in Papers from the Parasession on Functionalism, 276-336
- (1976), "Three Perspectives in the Functional Approach to Syntax", in Sound, Sign and Meaning. Quinquagenary of the Prague Linguistic Circle, edited by L. Matejka, p. 119-190
- Kuroda, S.Y. (1971), "Two Remarks on Pronominalization", Foundations of Language 7:2, 183-198
- Lakoff, G. (1968), "Pronouns and Reference". (Mimeographed by Indiana University Linguistics Club, Bloomington)
- (1970), "Repartee, or a Reply to "Negation, Conjunction and Quantifiers"", Foundations of Language 6, 389-422
- Langacker, R.W. (1969), "On Pronominalization and the Chain of Command", in Reibel and Schane (eds), Modern Studies in English, 160-186
- Lees, R.B. and Klima, E.S. (1963), "Rules for English Pronominalization", Language 39:1, 17-28
- MacColl, S. (1972), "A Note on Dougherty's Anapornia", Foundations of Language 9:1, 114-118
- McCawley, J.D. (1968), "Lexical Insertions in a Transformational Grammar without Deep Structure", in Darden et al., Papers from the Fourth Meeting of the Chicago Linguistic Society, 71-80
- (1970), "Where Do Noun Phrases come From?", in Jacobs and Rosenbaum (eds), Readings in Transformational Grammar, 166-183
- Pala, K. (1974), "Semantic Classes of Verbs and FSP", in Daneš 1974, 196-207

- Postal, P.M. (1971), Cross-Over Phenomena
- (1972a), "A Global Constraint on Pronominalization", Linguistic Inquiry 3:1, 35-59
- (1972b), "Some Further Limitations of Interpretative Theories of Anaphora", Linguistic Inquiry 3:3, 349-371
- Ross, J.R. (1967), "On the Cyclical Nature of English Pronominalization", in To Honor Roman Jakobson, 1669-1682
- Spangler, W.E. (1970), "Locative Restraints on the English Reflexive", Word 26:1, 114-118
- Svoboda, A. (1968), "The Hierarchy of Communicative Units and Field as Illustrated by English Attributive Constructions", Brno Studies in English 7, 49-102
- (1974), "On Two Communicative Dynamisms", in Daneš 1974, 38-42
- Teleman, U. (1970), "On Noun Phrases with Identical Reference", in Proceedings of the Conference of General and Nordic Linguistics in Reykjavik 1969
- Wasow, T. (1973), "More Migs and Pilots", Foundations of Language 9:3, 297-305
- (1975), "Anaphoric Pronouns and Bound Variables", Language 51:2, 368-383.

Note 8 After having finished this paper, I could read the doctoral dissertation of T. Wasow ("Anaphoric Relations in English", unpublished, Massachusetts Institute of Technology, 1972), which contains some counter-examples to Rule A, mostly sentences with generic NPs. ('When they are hungry, tigers can be dangerous.', 'If he has an ugly wife, a man should find a mistress.' etc. - The male chauvinist pig is Postal ("On Coreferential Complement Subject Deletion", Linguistic Inquiry 1:4, 439-500), not me.) Such 'generic pronouns' can be perhaps explained as having 'antecedents in the situation' in the sense of our knowledge of the words used in generic

sense. As for Wasow's other examples with indefinite but specific 'postcedents', ('After Bill kissed her, a certain young lady blushed repeatedly.', 'That he was not elected upset a certain leading politician.' etc.), these are examples of the speaker's power to present a part of a sentence as 'known to the listener', even when it is probable that the knowledge is not shared by the listener. The speaker talks in indeterminate terms, but he has a definite lady/politician in his mind.

Note 9 In the paper mentioned in Note 8, Wasow has also shown that some pronominal rhemes proper standing before their 'antecedents', are possible in special contexts, e.g.: 'Was it after you robbed the bank, that Mery left town? No. After HE robbed the bank, John left town.' It is quite impossible in sentences blocked by Rule B and C (cf. (35) and (37)). Some more examples where Rule D cannot be cancelled by any context:

\*It was HE who killed Peter's wife.

\*It was HIM Peter's wife killed.

\*It was HIS wife who killed Peter. (Compare with 'It was his WIFE who killed Peter'.)

\*It was HIM Peter's wife killed. etc.

(See also sentences (78) - (86) in the main text.)



VERGLEICHENDE STUDIEN ZUR PROSODIE SCHWEDISCHER DIALEKTE

Eva Gårding

Vortrag am Phonetischen Institut, Hamburg, 3.11.76

Das Schwedische und überhaupt die skandinavischen Sprachen sind zu prosodischen Untersuchungen sehr geeignet. Ein Grund dafür ist, dass die Prosodie dieser Sprachen ausser Ausdrucksmitteln wie Betonung und Satzintonation auch zweierlei Betonungsarten von mehrsilbigen Wörtern hat - die sogenannten Akzente.

Ein schwedisches Wort hat also einen von zwei Akzenten, Akzent 1 oder Akzent 2. Als Beispiel gebe ich ein berühmtes minimales Paar, [änden] A1, [ändən] A2. Die Betonung liegt auf der ersten Silbe in beiden Wörtern, aber das prosodische Muster ist verschieden. Für den grössten Teil des Wortschatzes ist der Akzent mit der Wortstruktur verbunden. Man kann daher den Akzent von der Wortstruktur ableiten und umgekehrt. Wer ein zweisilbiges Hauptwort in bestimmter Form mit Akzent 1 in der Lautkette erkennt, z.B. änden, der weiss sofort, dass die unbestimmte Form einsilbisch sein muss, also änd mit der Bedeutung Wildente. Wenn dieselbe Lautfolge Akzent 2 hat - ändən -, dann kann man darauf schliessen, dass der Stamm zweisilbisch ist, ände, was Geist bedeutet.

Das Weglassen des Akzentunterschieds beeinträchtigt sehr wenig die Kommunikationsfähigkeit. Die Bedeutung ist nämlich im Zusammenhang in fast allen Fällen eindeutig. Die schwedisch-sprechende Bevölkerung in Finnland zum Beispiel kommt ganz gut ohne Akzente aus.

Die Akzentregeln sind also im allgemeinen von der Wortstruktur abhängig, wie aus dem angeführten Beispiel - einsilbiger Stamm - Akzent 1, zweisilbiger Stamm - Akzent 2, hervorgeht. Die meisten dieser Akzentregeln sind den Dialekten gemeinsam. Ein Wort, das in Südschweden Akzent 1 hat, hat auch in Norden Akzent 1, und so weiter.

Die Manifestationsformen der Akzente sind aber in den Dialekten sehr verschieden, wie aus Figur 1 hervorgeht. Sie zeigt schematisierte



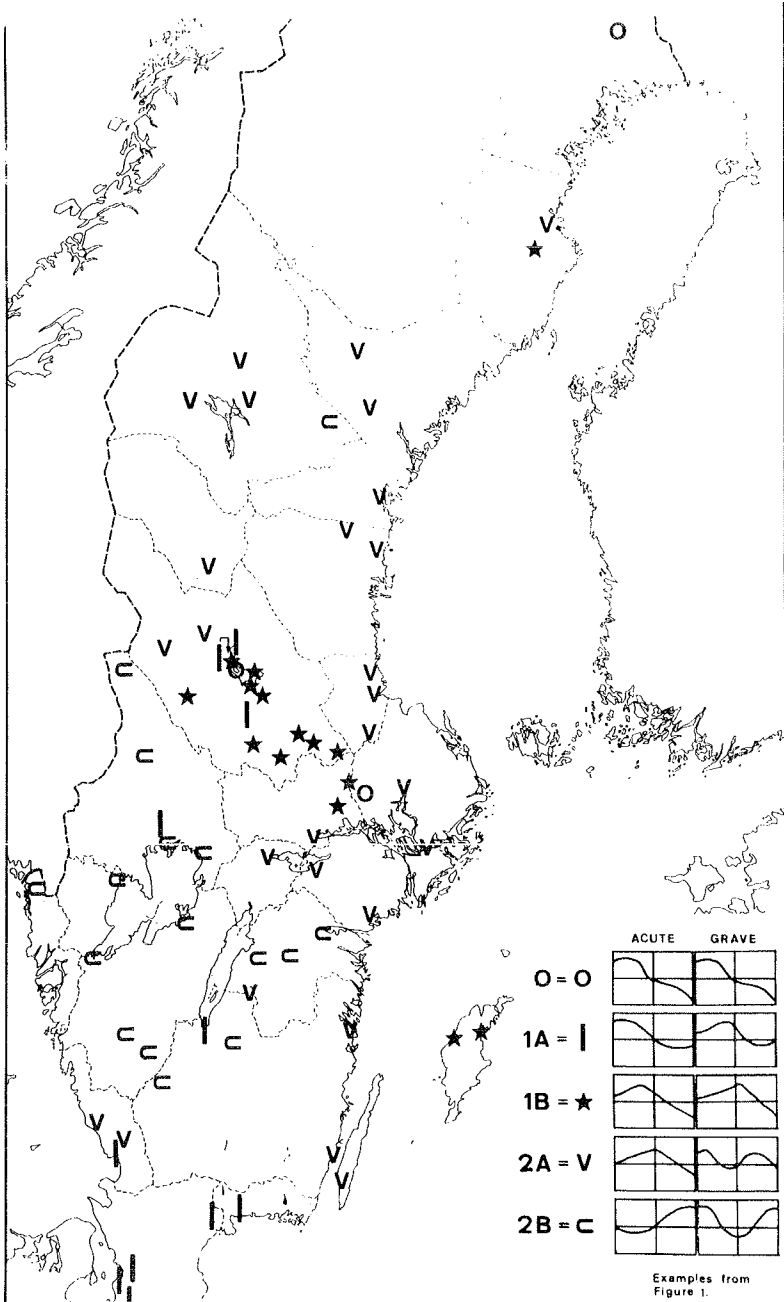


Fig. 2. Geographical distribution of accent types.



Grundtonskurven von Zweisilbern in neutralen Aussagesätzen in 100 Mundarten.

Dieses Material wurde von dem bekannten deutschen Phonetiker und Germanisten Ernst A. Meyer gesammelt.

Die Kurven oben links in der Figur stammen aus Stockholm. Die schematisierte Darstellung bezieht sich auf die beiden Silben in jedem Wort. Die senkrechte Linie in der Mitte bezeichnet die Silbengrenze, die waagerechte Linie die Zeitachse. Wir sehen hier, dass Akzent 1 von einem Tongipfel gekennzeichnet ist. Der Gipfel liegt etwa an der Silbengrenze. Akzent 2 hat zwei Tongipfel, einen für jede Silbe.

Im ersten Augenblick scheint der Variationsreichtum der Akzentformen ganz überwältigend zu sein. Aber es stellt sich heraus, dass sich diese Kurvenformen in zwei Hauptklassen gruppieren lassen. Für diese Gruppierung benutzen wir als Kriterien - wie Meyer - die Zahl der tonalen Wendepunkte und deren Lage im Verhältnis zu den Silben. Die eine Gruppe hat zweigipflige Tonkurven für Akzent 2, eingipflige für Akzent 1. Die andere Gruppe hat eingipflige Kurven für beide Akzente. Der Unterschied liegt hier in der Lage der Gipfel. Akzent 1 kommt immer früher als Akzent 2 oder mit Meyer's Worten: „Akzent 2 folgt immer in respektvollem Abstand auf Akzent 1“.

Figur 2 zeigt diese Analyse der Akzentformen und ihre geographische Verbreitung.

Beispiele der Akzenttypen enthält das Kästchen in der rechten Ecke.

Null: Kein Akzentunterschied

Typus 1: Ein Gipfel für Akzent 1 und einer für Akzent 2.

1A: Früh in der ersten Silbe für Akzent 1 und spät in derselben Silbe für Akzent 2.

1B: Spät in der ersten Silbe für Akzent 1 und früh in der Zweiten Silbe für Akzent 2.

Typus 2: Ein Gipfel für Akzent 1 und zwei für Akzent 2.

2A: Spät in der ersten Silbe für Akzent 1

2B: Ein Tal in der ersten Silbe und ein später Gipfel in der zweiten Silbe für Akzent 1.

Die geographische Verbreitung dieser Akzenttypen fällt mit in Schweden wohlbekannten Dialektgrenzen zusammen. Die alten Dialektgrenzen gründen sich auf lexikalische und morphologische Kriterien. Es gibt eingipflige Akzente 2 in Südschweden, das heisst in den alten dänischen Provinzen, in Dalarna in Mittelschweden und auf der Insel Gotland. In den übrigen Teilen Schwedens überwiegen die zweigipfligen Akzente 2. Fast alle schwedische Dialekte in Finnland haben keinen Akzentunterschied. Die Manifestationsformen der Akzente in den Dialekten geben der jeweiligen Intonation ihren besonderen Charakter. Wir haben sogar in einem Experiment gezeigt, dass typische Dialekte erkannt werden können, auch wenn die segmentale Information fehlt. Dazu haben wir den Kehlkopftton mit einem Laryngographen isoliert (Fourcin & Abberton 1971, Gårding et al. 1975).

Auf der Grundlage, die uns die Typologie gegeben hat, haben wir ein Projekt angefangen. Das Projekt heisst Schwedische Prosodie und wird vom Schwedischen Humanistischen Forschungsrat unterstützt. Es beschäftigt ausser mir selber auch Gösta Bruce, der seine Doktorarbeit in diesem Rahmen macht.

Unser Ziel ist es, ein Modell zu entwickeln, das Intonationskurven für alle schwedische Sätze in allen fünf Dialekttypen generiert, unabhängig vom Satztypus und von der Struktur des Satzes.

Die Methode besteht darin, Tonkurven von Phrasen mit Akzenten in verschiedenen Kontexten zu vergleichen. Auf diese Weise versuchen wir, die Beiträge von den Akzenten und den anderen Variablen zu zerlegen.

In diesem Vortrag werde ich mich auf drei Dialekte und auf ein begrenztes Material beschränken. Das Material besteht aus Nominalphrasen, die als neutrale Aussagen geäussert wurden. Die Testphrasen enthalten alle vier Kombinationen von Akzenten und eine variierende Zahl von unakzentuierten Silben zwischen den Akzenten. Die Mitteleiler wurden instruiert, drei verschiedene Betonungsmuster zu verwenden.

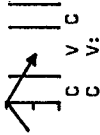
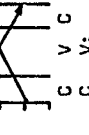

Stockholm (Central)	Skåne (South)	Helsinki (East)
Prep. rules	Delete A after P	Delete A after P
In a phrase with no P add P to the last A	In a phrase with no P add P to the last A	In a phrase with no P add P to the last A
High, neutral, low	High, neutral, low	High, neutral, low
A. Word accent commands	 A1	A
A2	 A2	 C V C C V:
P. Prominence command	rise after accent	replace accent fall by fall to low
S. Intonation commands (here statement)	1. fall after P-rise on the last syllable if last A has P if not, sharpen the fall of last A and delete rise 2. Beginning and end are neutral	replace accent fall by fall to low
Conventions	Fill in blank parts	fill in etc.
Input: Phrases with two accented words of which one is prominent.		
Example of input phrase: en långre numma [ en- <sup>A1</sup> lån- <sup>A2</sup> re-num-na ] Statement		

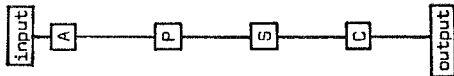
Fig. 3. Model and rules for pitch (f<sub>0</sub>).

Scheme

Central

South

East

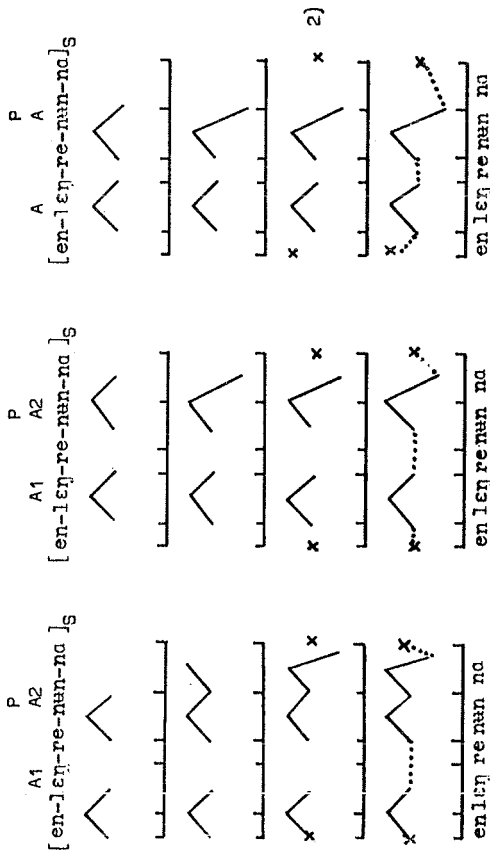


Rule 1. Apply A to accented syllables

2. Apply P to prominent syllable

3. Apply S

4. Apply C



1) When A is on the last syllable certain readjustments are necessary

2) x shows the application of S

Fig. 4. Application of pitch rules to noun phrases with two accented words (1) in three Swedish dialects.

In einem Muster wurde dem Adjektiv und dem Substantiv gleiches Gewicht, d.h. gleiche Betonung, gegeben oder/aber es wurde das Adjektiv oder das Substantiv betont.

Der jetzige Stand unseres Modells geht aus Figur 3 hervor. Das Modell kann die Tonkurve für eine Phrase generieren, die die folgende Information hat: Silbengrenzen, Phrasengrenzen, Wortakzente, Satzakkzent und Satztypus. Wir stellen uns vor, dass die Dauer der Segmente mit derselben Information schon berechnet worden ist. Eine typische Inputphrase ist unten in der Figur zu lesen: [en-1<sup>A1</sup>er-re-<sup>A2</sup>nanna] 'eine längere Nonne'. Die Regeln, oder Kommandos, repräsentieren die linguistischen Komponenten unseres Modells, Wortakzent, Satzakkzent und Satzintonation. Die Regeln schreiben Tonbewegungen zwischen 3 relativen Tonebenen vor.

Sehen wir uns zuerst die Kommandos für Stockholm an! Es gibt eine Steigung-Fall-Bewegung für beide Akzente, aber mit verschiedenen Domänen: Akzent 1 fängt früher an - schon in der präakkzentierten Silbe - und hört früher auf als Akzent 2.

Die Akzente für Südschwedisch werden auch durch eine Steigung - Fall-Bewegung realisiert und Akzent 2 kommt auch hier später als Akzent 1. Die Domänen sind aber im Verhältnis zum Mittelschwedischen verschoben. Akzent 1 markiert die akzentuierte Silbe und Akzent 2 umspannt auch die folgende unakzentuierte Silbe. Im Ostschwedischen gibt es, wie ich schon erwähnt habe, nur einen Akzent. Der ist dem südschwedischen Akzent 1 oder dem mittelschwedischen Akzent 2 ähnlich.

Der Satzakkzent - P steht für Prominenz - ist in Stockholm eine Steigung, die nach dem Akzentfall einsetzt. Infolge der Zeitverschiebung der Akzente trifft der Satzakkzent die akzentuierte Silbe nach Akzent 1 und die folgende unakzentuierte Silbe nach Akzent 2.

Der Satzakkzent in Südschweden und in Finnland ist derselbe, d.h. ein Fall zu der niedrigsten Ebene, der den Akzentfall ersetzt.

Die letzten Kommandos sind die Intonationskommandos der drei Dialekte. - S bedeutet Aussage (statement). Der Unterschied kann auf folgende Weise ausgedrückt werden. Der Ton ist in Stockholm durch den Satz-

akzent hinaufgezwungen worden und muss, um eine Aussage zu werden, auch wieder herunterkommen. Das geschieht im ersten Kommando. Die Bedingungen a) und b) betreffen den Zeitpunkt dieses Falles, der von der Struktur abhängig ist.

In den anderen Dialekten ist der Ton durch den Satzakzent schon auf der untersten Tonebene. Daher ist kein Fall mehr notwendig.

Mit dem zweiten Intonationskommando markieren wir die Ebenen, wo die Grundtonkurve anfängt und aufhört: Neutral für das Kontinental-schwedische, hoch für den Anfang der Kurve des Finnlandschwedischen.

Die letzte Komponente unseres Modells ist eine Ausfüllregel. Sie füllt die leeren Stellen der Kurve aus. Mit dieser Regel wird es möglich, eine beliebige Zahl von unakzentuierten Silben zwischen den Akzenten zu haben.

Die Regeln oben in Figur 3 sind abstrakte Distributionsregeln. Sie sind hier präparierende Regeln genannt. Die zweite Regel entspricht der berühmten Nuclear Stress Rule und ist gemeinsam für alle drei Dialekte. Diese Regel besagt, dass, wenn beide akzentuierte Wörter dasselbe Gewicht haben, der Satzakzent dem letzten Akzent zugewiesen werden soll. Die erste Regel ist den südschwedischen und ost-schwedischen Dialekten gemeinsam. Diese Regel verhindert die Tonbewegung des letzten Akzentes, ohne die Dauerverhältnisse zu verändern.

Figur 4 zeigt die Anwendung der Regeln. Als Input habe ich die Phrase en längre nunna gewählt. Der Satzakzent liegt auf dem letzten Wort. Das Attribut hat in diesem Beispiel Akzent 1, das Hauptwort Akzent 2. In der finnlandschwedischen Phrase gibt es nur akzentuierte Silben, keine Akzente. S in der Input-phrase bedeutet statement (Aussage).

Jetzt stellen wir uns vor, dass die Silben und Segmente die richtigen Dauerwerte erhalten haben. Dann kommen die Regeln. Regel 1 gibt die Akzentbewegungen. Rechts sehen Sie das Ergebnis der Regeln. Die akzentuierten Silben sind entsprechend markiert worden. Regel 2 gibt den Satzakzent - angehängt im Zentralschwedischen, akzentmodifizierend im Süd- und Ostschwedischen. Regel 3 gibt die Aussagekonfiguration

der Kurve, Schlussfall für Zentralschwedisch sowie Anfangs- und Schlusspunkte für alle Dialekte. Regel 4 schliesslich vollzieht die Kurven durch Interpolation - durch punktierte Linien markiert.

Die Schematisierung verbirgt einige interessante Unregelmässigkeiten. Eigentlich haben wir für diese schon einige Anpassungsregeln eingeführt, die ich aber in dieser Aufstellung weggelassen habe. Mit diesen Regeln bekommen unsere generierten Kurven eine bessere Anpassung an die beobachteten Daten.

Diese Anpassungsregeln spiegeln teils tonale Koartikulationseffekte wieder, teils eine Prioritierung von Kommandos. Eine solche Prioritierung muss zustande kommen, wenn der Sprecher nicht Zeit hat, alle Kommandos auszuführen, z.B. in einem einsilbigen Satz. In diesem Fall wird der Anfang des Akzents zum Wort-Anfang hinverschoben, um dem Satzakzent und der Satzintonation genügend Platz zu geben.

Mit unserem Modell und unserer Methode glauben wir, dass wir ein klareres Bild von den prosodischen Verhältnissen in den Dialekten erhalten werden.

Auch ein flüchtiger Blick auf Figur 3 zeigt, dass die Prosodie von Osten, d.h. Finnland, und Süden ziemlich gleich ist. In beiden Dialekten wird das akzentuierte Wort als Steigung - Fall realisiert. Der Satzakzent modifiziert auf dieselbe Weise die Konfiguration des Falles. Der Unterschied liegt in der Zahl der Akzentkommando. Auch die Distributionsregeln sind gleich.

Wenn wir einen Augenblick zu unserer ursprünglichen Typologie in Figur 1 zurückkehren, dann sehen wir, dass wir den sogenannten zweigipfligen Akzenten 2 eine mehr adequate Analyse geben können. Diese Akzente bestehen aus einem Akzentteil - das erste Dreieck - einem Satzakzent, die Steigung, und einem Intonationsteil, dem Schlussfall. Es wurde früher behauptet, dass nur Akzent 2 aus einem Akzentteil bestände, der eingipflige Akzent 1 wäre die Manifestation von Satzakzent und Satzintonation (vgl. die Übersicht in Gårding 1973). Gösta Bruce's Analyse von längeren Phrasen mit Akzenten in verschiedenen Positionen hat aber gezeigt, dass auch Akzent 1 ein konstantes Tonmerkmal hat (Bruce 1975).

Aber die Typologie besteht, obwohl es vielleicht richtiger wäre, von einer prosodischen Typologie zu sprechen. Es gibt noch immer zwei Hauptgruppen von Dialekten, die in der akustischen Registrierung und auch für das Ohr sehr verschieden sind.

Tatsächlich sind die Dialekte mit eingipfligem Akzent 2 ziemlich ähnlich anderen germanischen Sprachen wie Deutsch und English und könnten vielleicht als eine prosodische Übergangszone angesehen werden.

#### LITERATURVERZEICHNIS

- Bruce G. 1975. Stockholm accents in focus. Working Papers 10. Phonetics Laboratory, Universität, Lund
- Bruce G. 1975. Swedish accents in sentence perspective. Paper read to the 8th Congress of Phonetic Sciences. Siehe kommende Abhandlung: in Travaux de l'institut de linguistique de Lund
- Fourcin A.J. und E. Abberton. 1971. First application of a new laryngograph. Medical and Biological Illustration 21
- Meyer E.A. 1937. Die Intonation im Schwedischen, I, Die Sveamundarten. Studies Scand. Philol. No. 10. Universität, Stockholm
- Meyer E.A. 1954. Die Intonation im Schwedischen, II, Die norrländischen Mundarten. Wie oben. No. 11
- Gårding E. 1970. Word tones and larynx muscles. Working Papers 3. Phonetics Laboratory, Universität, Lund
- Gårding E. 1973. The Scandinavian Word Accents. Working Papers 8. Phonetics Laboratory, Universität, Lund
- Gårding E. 1975. Toward a prosodic typology for Swedish dialects. Proc. 2nd Congress of Nordic and General linguistics. Umeå
- Gårding E. 1976. The importance of Turning points for the pitch patterns of Swedish accents. Symposium on stress and accent. University of Southern California. Feb. 28-29
- Gårding E. und P. Lindblad. 1973. Constancy and variation in Swedish word accent patterns. Working Papers 7. Phonetics Laboratory, Universität, Lund
- Gårding E., R. Bannert, A.-Chr. Bredvad-Jensen und K. Nauclér. 1974.



Talar skåningarna svenska? In Platzack (Hrsg.) Förhandlingar för svenskans beskrivning 8

Gårding E., O. Fujimura, H. Hirose und Z. Simada. 1975. Laryngeal control of Swedish word accents. Working Papers 10. Phonetics Laboratory, Universitätt, Lund

L'aspiration des occlusives après [s].

Magnús Pétursson

Introduction.

C'est un trait connu des langues germaniques que les consonnes p t k qui sont généralement aspirées en début de mot apparaissent sans aspiration si elles sont précédées d'un [s] ou d'une autre constrictive sourde. L'aspiration est pourtant conservée, même après [s], si une limite de mot ou de morphème se place immédiatement après le [s], p. ex.:

	<u>aspirée</u>	<u>non aspirée</u>	<u>aspirée après [s]</u>
Anglais:	peak	speak	this peak
Allemand:	Paar	sparen [ʃ-]	das Paar [s-]
	Tat	starr [ʃ-]	Misst <u>at</u> [-s-]
Danois:	på	spare	pas på
Islandais:	pinni	spinna	laus pinni
	tak	stakur	Íst <u>ak</u>

Après [s] l'aspiration apparaît donc dans les langues germaniques modernes comme trait démarcatif aphonématique ou comme jointure à la limite de mot ou de morphème (Lehiste 1960, p. 26-28; Trubetzkoy 1962, p. 242, 255-256). Il n'en est pas ainsi dans toutes les langues. Dans de nombreuses langues indiennes, comme p. ex. malayalam, hindi et sindhi <sup>1</sup> (Doi, 1974), l'aspiration peut apparaître après [s] dans le même morphème. La particularité des langues germaniques de n'accepter l'aspiration après [s] que si les deux consonnes appartiennent à des mots ou à des morphèmes différents, n'est donc pas un trait universel (Lindqvist 1972, p. 20). C'est plutôt un trait caractéristique des langues germaniques, bien qu'il n'ait que rarement été reconnu comme tel.

1. Nous remercions M. le professeur Paroo Nihalani, Hyderabad (Inde), pour les précieuses informations qu'il a bien voulu nous donner à propos de l'aspiration après s dans de nombreuses langues de l'Inde.

Du point de vue articulatoire il se pose la question suivante: Comment la glotte se comporte-t-elle pour réaliser l'aspiration après s? Dans les langues germaniques cette question peut seulement être examinée si on recourt à des combinaisons de mots ou à des mots composés. Le comportement de la glotte pour les groupes sp-, st-, sk- est connu (Pétursson 1975, p. 115-120; Lindqvist 1972) et des hypothèses ont été avancées sur le comportement de la glotte pour réaliser l'aspiration après s ou après une autre constrictive sourde (Pétursson, sous presse).

Le but de cette modeste contribution est l'examen du comportement de la glotte à la limite de mot et de morphème dans une langue germanique, l'islandais moderne. En même temps nous espérons attirer l'attention des phonéticiens sur un trait commun aux langues germaniques (à l'exception du néerlandais qui n'a pas d'occlusives aspirées) sur lequel ils n'ont presque pas porté leur attention (Lisker et al. 1969, p. 1545 ont délibérément choisi de ne pas aborder le problème des occlusives après s à cause du statut linguistique ambigu de ces occlusives).

#### Méthodes expérimentales et matériaux linguistiques.

Nous avons examiné l'activité glottale avec le glottographe photoélectrique Frøkjær-Jensen (Pétursson 1975, p. 109-110, 123; Löfqvist/Pétursson, in press)<sup>2</sup>. Simultanément nous avons enregistré l'oscillogramme et le débit d'air. La disposition des appareils est celle que nous avons décrite dans notre travail récent où nous avons également discuté les problèmes de l'interprétation des glottogrammes (Pétursson 1975, p. 109-110; Sawashima 1974).

Les matériaux linguistiques sont des mots et des combinaisons de mots islandais qui ont été lus isolément, précédés par le mot-clé segðu "dis", impératif du verbe segja "dire". Les mots utilisés pour cet enregistrement sont les suivants:

2. Les enregistrements ont été réalisés à l'Institut de Phonétique de Lund le 11 juin 1976. Nous remercions tout spécialement M. Anders Löfqvist et Mme Eva Gårding pour leur aide et encouragement.

1. Aspiration après [s] à la limite de deux mots:

[-s #t <sup>h</sup> -]	<u>laus tǒng</u> [lǒi:s t <sup>h</sup> ǒiŋk]	"tenaille non attachée"
	<u>laus tími</u> [lǒi:s t <sup>h</sup> i:mI]	"temps libre"
	<u>laus tǒk</u> [lǒi:s t <sup>h</sup> ǒ:k]	"faible prise"
[-s #c <sup>h</sup> -]	<u>laus kefli</u> [lǒi:s c <sup>h</sup> eplI]	"cylindres non attachés"
	<u>laus kind</u> [lǒi:s c <sup>h</sup> Int]	"mouton non attaché"
[-s #p <sup>h</sup> -]	<u>laus pinni</u> [lǒi:s p <sup>h</sup> In:I]	"bâton non attaché"
	<u>laus pund</u> [lǒi:s p <sup>h</sup> Ynt]	"livres non attachées" (livres=unités de poids)

2. Absence d'aspiration après [s] dans le même mot:

[st-]	<u>stök</u> [stǒ:k]	"isolée", fém. de <u>stakur</u>
	<u>stími</u> [sti:mI]	"qu'il avance" subj. prés. de <u>stíma</u>
[sc-]	<u>skefli</u> [sceplI]	"qu'il se forme une congère", subj. prés. de <u>skefla</u>
	(í) <u>skyndi</u> [scIntI]	"rapidement"
[sp-]	<u>spinni</u> [spIn:I]	"qu'il file", subj. prés. de <u>spinna</u>
	<u>spunnum</u> [spYn:Ym]	"nous avons filé", passé de <u>spinna</u>

3. Aspiration après [s] à la limite de deux morphèmes:

[-st <sup>h</sup> -]	<u>ístǒng</u> [i:st <sup>h</sup> ǒiŋk]	"tenaille pour prendre de la glace"
	<u>ístak</u> [i:st <sup>h</sup> a:k]	(nom propre)

Chaque mot a été lu 20 fois avec débit normal. En outre les mots du premier groupe ont été lus cinq fois avec débit rapide pour voir si le comportement de la glotte changeait en fonction de la vitesse de lecture. Les paramètres suivants ont été mesurés:

1) durée de la consonne [s]; 2) durée de l'occlusion; 3) durée de l'aspiration; 4) durée de l'explosion des occlusives non aspirées (deuxième groupe); 5) le moment de l'ouverture glottale maximale à l'intérieur du [s]; 6) le moment de l'ouverture glottale maximale dans la consonne aspirée.

Le moment de l'ouverture glottale maximale est mesuré à partir du début de la consonne. La durée est mesurée en milli-secondes.

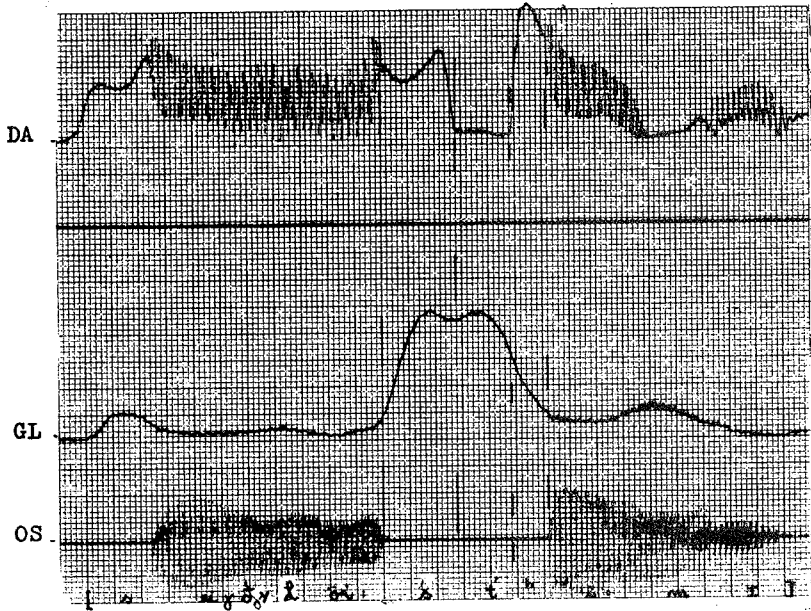


Fig. 1. Glottogramme des mots segou laus timi "dis: temps libre".  
 DA = débit d'air; GL = photoglottogramme; OS = oscillogramme. On peut observer deux maxima d'ouverture glottale, chaque maximum correspondant aux consonnes [s] et [t<sup>h</sup>] respectivement.

#### Résultats de l'analyse.

Les résultats des mesures sont résumés dans le tableau 1. Il ressort du tableau 1 que la durée de la consonne [s] est plus brève lorsque le [s] se trouve à la fin de syllabe devant une consonne aspirée dans les séquences [-s # t<sup>h</sup>-, -s # c<sup>h</sup>-, -s # p<sup>h</sup>-] que si le [s] se trouve au début de syllabe devant une occlusive non aspirée dans les groupes [st-, sc-, sp-]. Le groupe [-st<sup>h</sup>-] à la limite de deux morphèmes occupe une position intermédiaire. On peut expliquer les différences de durée observées par le fait que dans les groupes [st-, sc-, sp-] la consonne [s] se trouve dans la position explosive de la syllabe, alors que le [s] à la limite de deux mots se trouve en position implusive. La consonne implusive est généralement moins énergique que la consonne explosive et aura par conséquent normalement une durée moindre que celle-ci.

Tableau 1. Résultat des mesures.

Groupe	Nombre d'ex.	1	2	3	4	5	6	Mots
-s#t <sup>h</sup> -	20	123	90	33		64	73	laus <u>t</u> öng
	20	117	106	43		72	49	laus <u>t</u> ími
	20	118	91	45		77	60	laus <u>t</u> ök
Moyenne		119	95	40,3		71	60,3	
-s#c <sup>h</sup> -	20	100	72	54		95	31	laus <u>k</u> efli
	20	116	89	51		95	44	laus <u>k</u> ind
Moyenne		108	80	52,5		95	37,5	
-s#p <sup>h</sup> -	20	112	115	33		68	85	laus <u>p</u> inni
	20	113	107	39		79	85	laus <u>p</u> und
Moyenne		112,5	111	36		73	85	
st-	20	146	97		12	61		<u>st</u> ök
	20	149	98		15	60		<u>st</u> ími
Moyenne		147,5	97,5		13,5	60,5		
sc-	20	126	79		12	68		<u>s</u> kefli
	20	128	73		18	76		(í) <u>s</u> kyndi
Moyenne		127	76		15	72		
sp-	20	168	95		12	62		<u>s</u> pinni
	20	127	91		16	62		<u>s</u> punnum
Moyenne		147,5	93		14	62		
-st <sup>h</sup> -	19	147	86	34		76	53	í <u>st</u> öng
	19	126	86	32		62	51	í <u>st</u> ak
Moyenne		136,5	86	33		69	52,5	

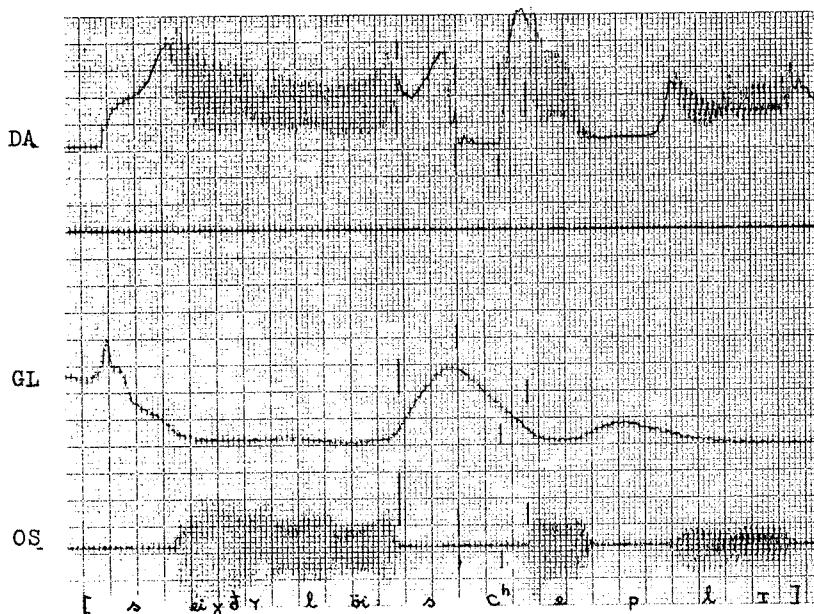


Fig. 2. Glottogramme des mots segou laus kefli "dis: cylindres non attachés". Un seul maximum d'ouverture glottale est réalisé vers la fin de la consonne [s].

Bien que ces différences soient ainsi en quelque sorte automatiquement réglées, il est néanmoins possible qu'elles puissent constituer un signal démarcatif pour identifier le début de mot ou de morphème.

Si l'on compare le moment du maximum de l'ouverture glottale à l'intérieur du [s] dans les groupes [-s#t<sup>h</sup>-, -s#c<sup>h</sup>-, -s#p<sup>h</sup>-] d'une part et dans les groupes [st-, sc-, sp-] d'autre part on peut voir que le maximum d'ouverture intervient en général plus tôt dans les groupes [st-, sc-, sp-]. On peut penser que cette localisation du maximum d'ouverture glottale est un moyen pour assurer que l'occlusive subséquente reste sans aspiration. Si l'ouverture glottale maximale a lieu tôt dans la consonne [s], la glotte aura un temps suffisamment long pour se fermer avant la rupture de l'occlusion. Au niveau glottal ces groupes sont traités comme une unité avec un seul maximum d'ouverture pour les deux consonnes (fig. 5).

Il en est tout autrement avec les groupes dans lesquels une limite de mot ou de morphème sépare les deux consonnes et où

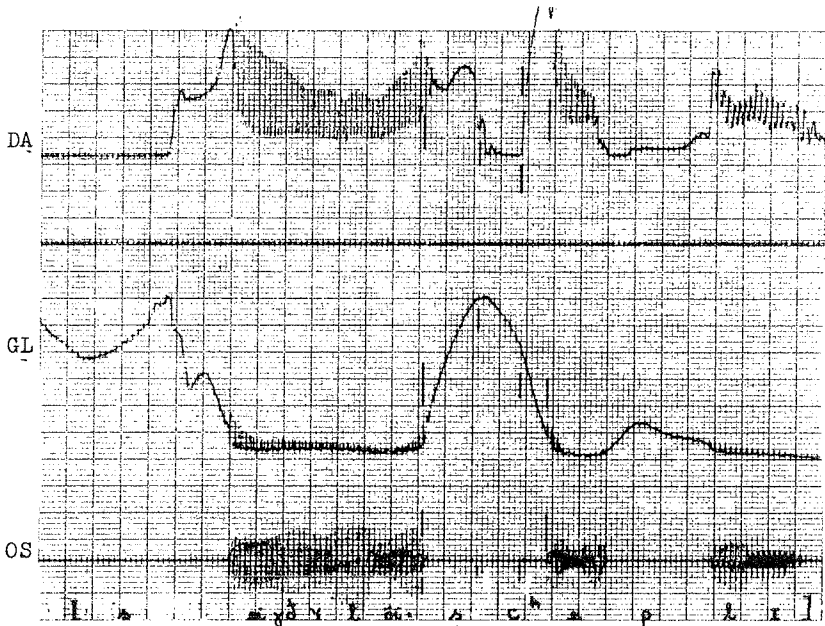


Fig. 3. Glottogramme des mots segðu laus kefli "dis: cylindres non attachés". Un seul maximum d'ouverture glottale est réalisé au début de la deuxième consonne [c<sup>h</sup>].

l'occlusive subséquente est aspirée. Dans ce cas on peut observer trois modèles de coordination de l'ouverture glottale maximale avec l'activité des organes supraglottaux (fig. 7):

1. Le cas le plus fréquent est qu'il y ait deux maxima d'ouverture glottale, chaque maximum correspondant à une consonne (fig. 1). Très fréquemment le deuxième sommet d'ouverture glottale reste moins prononcé que le premier (fig. 4; Frøkjær-Jensen et al. 1971, fig. 5 p. 138).
2. Un cas moins fréquent est que le maximum d'ouverture glottale intervienne tard à l'intérieur de la consonne [s] et que la glotte se ferme assez lentement pour que la consonne subséquente puisse être aspirée.
3. Le troisième cas est la réalisation d'un seul maximum d'ouverture glottale à l'intérieur de l'occlusive subséquente (fig. 3). La glotte se ferme ensuite lentement et l'occlusive reste aspirée.



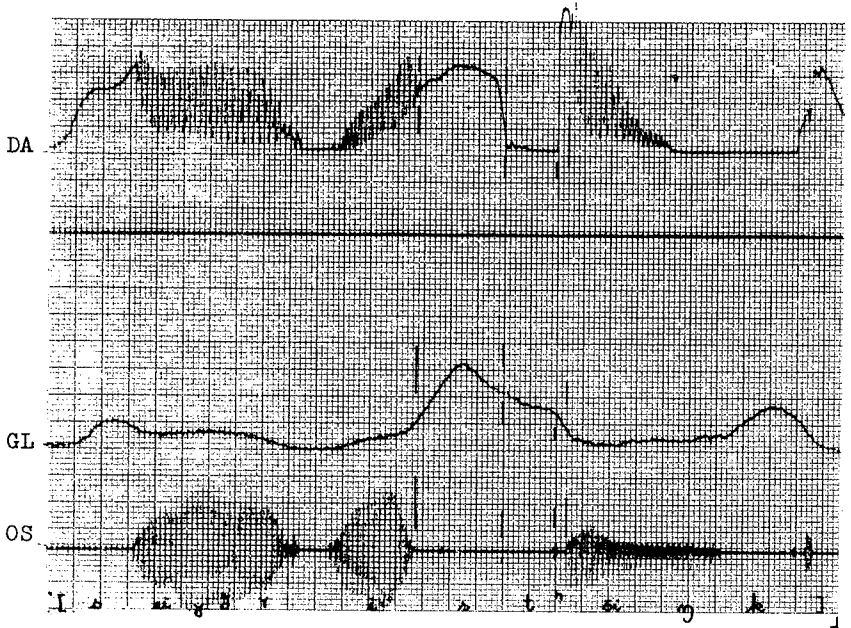


Fig. 4. Glottogramme des mots segðu ístöng "dis: tenaille pour prendre de la glace". Il y a deux maxima d'ouverture glottale, l'un à l'intérieur de la consonne [s], l'autre moins prononcé 10 msec. avant l'explosion du [t<sup>h</sup>].

Il est à noter que dans tous les cas la glotte reste grande ouverte durant la consonne [s]. Nous n'avons pas observé une petite ouverture glottale durant la prononciation du [s], comme on aurait pu le prévoir théoriquement (Pétursson, sous presse).

#### Discussion.

D'après les données examinées il semble évident que les mouvements d'ouverture et de fermeture de la glotte sont déterminés en fonction des unités linguistiques dont les mots respectifs sont formés. S'il n'en était pas ainsi et si un autre principe réglait les mouvements glottaux, p. ex. une contrainte physiologique ou aérodynamique, il est probable qu'on pourrait observer un modèle unique pour les groupes s+occlusive aspirée. Or, ceci n'est pas le cas.

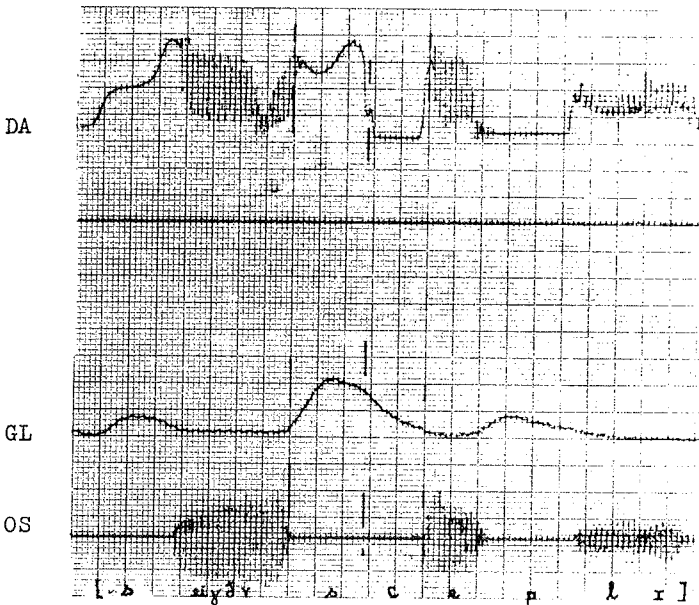


Fig. 5. Glottogramme des mots segðu skefli "dis: qu'il se forme une congère". Il se forme un seul maximum d'ouverture glottale pour les deux consonnes. Ce maximum se situe toujours à l'intérieur de la consonne [s].

Les trois modèles observés suggèrent un autre principe, à savoir les unités linguistiques dont les mots sont formés (fig. 7). Les deux maxima d'ouverture glottale paraissent également suggérer qu'il y a une innervation séparée pour chacune des consonnes. Quant aux exemples où l'on observe un seul maximum d'ouverture glottale ils permettent également de conclure que l'innervation est réalisée séparément pour chaque consonne, car la localisation du maximum est variable et elle est telle que l'aspiration peut être réalisée.

Pour les groupes [st-, sc-, sp-] il semble au contraire qu'au niveau glottal le groupe entier est traité comme une unité. Pour ces groupes il y a un seul modèle observé, à savoir que le maximum d'ouverture glottale intervienne dans le [s], généralement assez tôt, en tout cas dans les deux premiers tiers de la durée de la consonne.

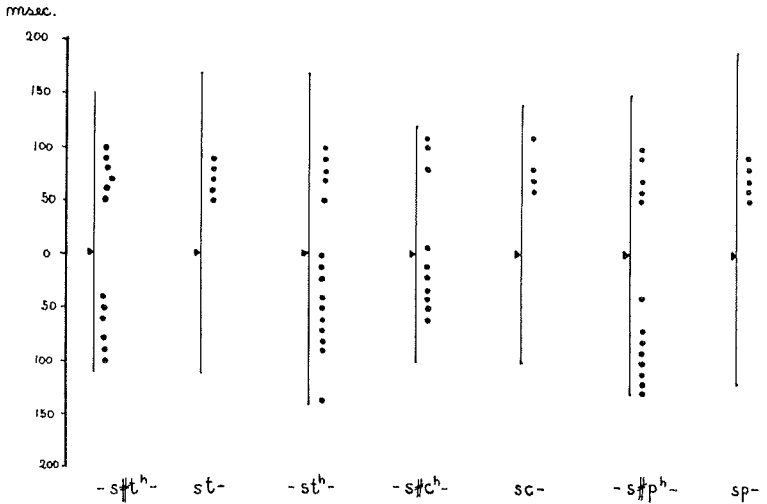


Fig. 6. Localisation des maxima d'ouverture glottale indiquée par des points noirs par rapport à la durée des consonnes constituant les groupes.  $\blacktriangleright$  indique la limite entre les deux consonnes et la ligne verticale indique la durée maximale mesurée pour les exemples enregistrés.

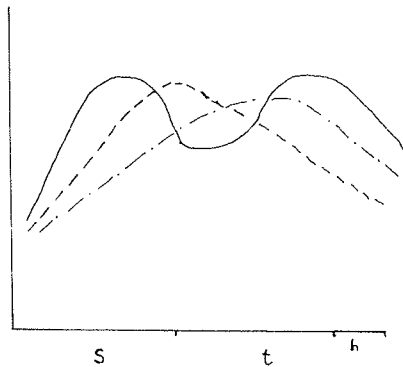


Fig. 7. Schéma du comportement glottal dans la prononciation des groupes s+occlusive aspirée. L'ouverture glottale est indiquée en abscisse. — deux maxima d'ouverture, chaque maximum correspondant à une consonne; --- un seul maximum d'ouverture dans la consonne [s]; -·-·- un seul maximum d'ouverture dans l'occlusive aspirée.

Ceci ressort de façon particulièrement claire de la fig. 6 qui représente la distribution des maxima d'ouverture glottale dans les groupes examinés par rapport à la durée des consonnes.

Comme la fig. 6 le montre, la distribution des maxima d'ouverture glottale est pratiquement la même pour tous les groupes en ce qui concerne la consonne [s]. Le maximum d'ouverture est généralement localisé au milieu de la durée du [s]. Uniquement en cas du groupe [-s#c<sup>h</sup>-] nous avons pu observer que le maximum d'ouverture glottale puisse se trouver vers la fin du [s] (figs. 2 et 6), mais dans ce cas il y a un seul maximum d'ouverture pour les deux consonnes.

La distribution des maxima d'ouverture glottale est plus large pour les consonnes aspirées que pour le [s]. S'il y a deux maxima d'ouverture pour les groupes de s+occlusive aspirée, le deuxième maximum d'ouverture est généralement localisé assez tard dans la phase occlusive. Il peut même être localisé au moment de la rupture de l'occlusion. Nous n'avons pas observé des cas où il soit localisé après la rupture de l'occlusion comme cela est souvent le cas pour les occlusives danoises. S'il y a un seul maximum d'ouverture glottale pour les deux consonnes, il peut être localisé très tôt dans la phase occlusive, même au début de la consonne. Le groupe [-st<sup>h</sup>-] a la distribution des maxima la plus large (fig. 6). Pour ce groupe le maximum peut être localisé à pratiquement n'importe quel moment de l'occlusion.

Les trois modèles de coordination du mouvement glottal (fig. 7) observés pour les groupes de s+occlusive aspirée ne semblent pas être réalisés en fonction du débit. Dans les exemples lus avec débit rapide les trois modèles ont été observés sans qu'il soit possible de découvrir un facteur régulier qui déciderait de l'apparition de chacun d'eux. Frøkjær-Jensen et al. (1971, p. 138) semblent suggérer que le débit pourrait altérer la forme de l'ouverture glottale, mais nos enregistrements ne contiennent pas d'exemples qui confirmeraient cette hypothèse.

En résumant: ce qui nous semble la conclusion la plus importante de cette étude c'est que de toute évidence les mouvements glottaux sont gouvernés par des unités linguistiques sous-jacentes et non pas par un autre principe comme p. ex.

le principe de l'économie des mouvements glottaux. Si le principe de l'économie gouvernait les mouvements de la glotte, il est difficilement concevable qu'on réaliserait une aspiration après une constrictive sourde, car il est certainement plus économique de réaliser un seul maximum d'ouverture comme pour les groupes [st-, sc-, sp-]. C'est d'ailleurs le modèle à un maximum qui semble prédominer dans toutes les langues pour toutes les articulations (comp. Uchita 1974, p. 711 pour le japonais). Pour les groupes de s+occlusive aspirée ceci n'est pourtant pas le cas. Dès que les unités et les oppositions linguistiques l'exigent, l'aspiration peut apparaître là où il est nécessaire pour assurer une distinction linguistique, même si du point de vue articulatoire cela suppose une violation du principe de l'économie des mouvements, principe qui — personne n'en doute — joue pourtant un rôle important dans l'aspect articulatoire du langage.

Phonetisches Institut  
der Universität Hamburg

Bibliographie.

- Doi, K. (1974): [On Hindi aspirates] (en japonais avec résumé en anglais); dans: World Papers in Phonetics, Festschrift for Dr. Onishi's Kiju, pp. 725-730 (The Phonetic Society of Japan, Tokyo).
- Frøkjær-Jensen, B.; Ludvigsen, C.; and Rischel, J. (1971): A glottographic study of some Danish consonants; dans: Form and Substance, Phonetic and Linguistic Papers presented to Eli Fischer-Jørgensen, pp. 123-140 (Akademisk Forlag, København).
- Lehiste, I. (1960): An acoustic-phonetic study of internal open juncture, Phonetica, Supplementum ad vol 5. (Karger, Basel).
- Lindqvist, J. (1972): Laryngeal articulation studied on Swedish subjects, STL-QFSR 2-3/1972, pp. 10-27.

- Lisker, L.; Abramson, A.S.; Cooper, F.S. and Schvey, M.H. (1969):  
Transillumination of the larynx in running speech, Journal  
Acoust. Soc. America 45, 1544-1546.
- Löfqvist, A. and Pétursson, M. (in press): Swedish and Icelandic  
stops. A glottographic study.
- Pétursson, M. (1975): Étude glottographique de quelques con-  
sonnes islandaises, Working Papers 12, Phonetics Labora-  
tory, Lund University, pp. 107-130.
- Pétursson, M. (sous presse): Aspiration et activité glottale.  
Examen expérimental à partir de consonnes islandaises.
- Sawashima, M. (1974): Laryngeal research in experimental  
phonetics; dans: T.A. Seboek (éd.): Current Trends in  
Linguistics, vol. 12, 4 pp. 2303-2348 (Mouton & Co.,  
The Hague).
- Trubetzkoy, N. S. (1962): Grundzüge der Phonologie (Vanden-  
hoek & Ruprecht, Göttingen, 3. Aufl.).
- Uchita, S. (1974): Glottographisch-morphologische Untersuchung  
an Konsonanten des Japanischen; dans: World Papers in  
Phonetics, Festschrift for Dr. Onishi's Kiju, pp. 709-718  
(The Phonetic Society of Japan, Tokyo).



IDENTIFICATION TEST CONCERNING ISOLATED DISYLLABIC SWEDISH  
ACCENT 1 AND ACCENT 2 WORDS

Hans Randler

INTRODUCTION

A characteristic feature of the prosody of Swedish (and other Scandinavian languages) is the word accent (word tone, toneme), giving an opposition between words with acute accent (accent 1) and grave accent (accent 2). The opposition is limited to polysyllabic words. The distribution rules are fairly uniform in all Scandinavian languages. In generative descriptions of Norwegian and Swedish accent 2 is considered to be marked and accent 1 is treated as stress. From a communicative viewpoint the constructive function of accent 2 is the most important, as it signals connection between syllables within the same word (Elert, 1970).

The word tones are manifested phonetically as a difference in the fundamental frequency curve of the syllable sequence. Perceptually this is the most important cue for word accent distinction (Malmberg, 1955). In minimal pairs the tone curve is the only distinctive factor within one and the same dialect (e.g. anden, "the duck", accent 1 and anden, "the spirit", accent 2).

The accent opposition exists with a few exceptions in all Swedish dialects. The phonetic realizations of the accents display a great variety of pitch patterns for different dialects. This has been shown in several investigations. On the basis of Meyer's (1937, 1954) material of bisyllables in statement intonation Gårding and Lindblad (1973) set up a tentative tonal typology for Swedish dialects (Table 1). Many listening tests have been carried out to test the intelligibility of the word accents in isolated words within and between dialects (Hadding 1961; Johansson 1970; Fintoft 1970). In all these investigations the listening groups made the highest identification score when they listened to their own dialect, which is taken



to confirm the assumption that a listener's own fundamental frequency patterns are perceptually important.

The different dialect manifestations of the word accents in Table 1 have also been reported by other researchers than Meyer. There are several collections of accent data from southern Swedish (1A) and Stockholm Swedish (2A), e.g. Malmberg (1955), Hadding (1961), Gårding and Lindblad (1973) (Fig. 1). The available accent data from the island of Gotland (1B) (Klintberg 1884; Pettersson 1953; Gårding and Lindblad 1973) shows one peak each for the two accents. Klintberg's descriptions were based on purely auditory observation. (Cf. also below.)

#### THE PRESENT INVESTIGATION

##### Purpose

The present investigation uses listeners speaking the dialect of Gotland (1B) in a listening test with isolated disyllabic accent 1 and accent 2 words. The purpose of the investigation is to test the validity of earlier results in similar tests and to try to answer a few questions regarding the typology mentioned above, where Swedish dialects are divided into two main groups according to the pitch curve of accent 2.

As mentioned, the listener seems to be dependent on his own fundamental frequency pattern when identifying isolated accent 1 and 2 words. We cannot interpret acoustic data in perceptual terms (there is no one-to-one correspondence between the levels) but Table 1 raises several interesting questions. The dialects of Skåne and Gotland have pitch curves of a similar type, i.e. one peak in each accent. How will that affect the listener's identification ability? According to Table 1 the position of the pitch peak is the same in accent 2 of 1A (Skåne) as in accent 1 of 1B (Gotland). What effect will this resemblance have on the results of a listening test? What is the perceptual relationship like between the dialect of Gotland and those in the other groups? Both Hadding (1961) and Johansson (1970)

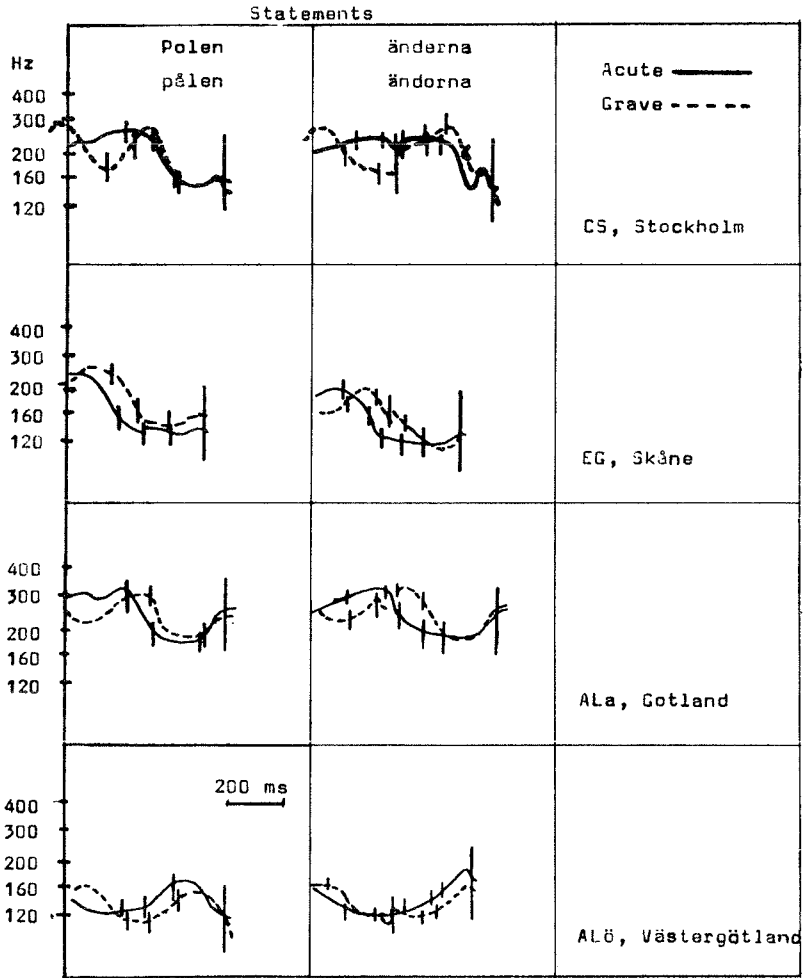


Fig. 1. Dialect variation of accents in statements.  
From Gårding and Lindblad (1973).

	Accent 1	Accent 2	Region
Type 0	One peak In the stressed syllable	One peak In the stressed syllable	Isolated marginal cases in the far north and in Finland
	No distinctive difference		
Type 1	One peak	One peak	
Type 1A	Early in the stressed syllable	Late in the stressed syllable	South (former Danish provinces)
Type 1B	Late in the stressed syllable	Early in the post- tonic syllable	Gotland (in the Baltic), Dalecarlia
Type 2	One peak	Two peaks	
Type 2A	In the stressed syllable	One in each syllable	Central Sweden (Svea dialects)
Type 2B	In the post-tonic syllable	One in each syllable	Between southern and central Sweden (Göta dialects)

Table 1. Tentative tonal typology for accent in Swedish dialects.  
From Gårding and Lindblad (1973).

reported more correct responses to the accent 2 words. Will the proportions be similar in this test? If so - how can this be explained? Furthermore, are there any perceptual differences between listeners from different parts of the island? Does recognition of a dialect improve a listener's ability to identify isolated accent 1 and accent 2 words?

#### The test words

A listening test was designed with accent 1 and accent 2 words (minimal pairs) from the four dialect areas mentioned above (Skåne, Gotland, Stockholm and Västergötland). The words were presented in random order and the listeners were asked to give the meaning of each word (forced choice). Essentially the same method was used by Kloster-Jensen (1961), Hadding (1961), Johansson (1970) and Fintoft (1970).

Considering the purpose and procedure of the investigation the selection of disyllabic test words was made with the following points in mind:

1. Their segmental differences should not vary too much between the dialects<sup>1</sup>.
2. The minimal pairs selected should have different vowel manifestations in the stressed syllables.
3. The meaning of the words should be unambiguous and easily understood.
4. The words should be natural and common to the listeners.
5. It must be possible to visualize the meaning of each word by drawing a simple picture.

The following pairs were chosen:

accent 1	accent 2
<u>˘vaken</u> "the hole in the ice"	<u>˘vaken</u> "awake"
<u>˘skotten</u> "the shots"	<u>˘skotten</u> "the Scot"
<u>˘anden</u> "the duck"	<u>˘anden</u> "the spirit"

The test words were read in the sentence frame jag såg ... där (I saw ... there)<sup>2</sup>.

The relationship between the word tones and the sentence intonation has been studied by Hadding (1961). She found that accent distinctions are preserved in different intonation patterns, even in rapid speech. In a comment on Fintoft (1970), Witting(1972, p. 41) says that the sentence intonation may affect the accent of a word in final position of an utterance. He thinks it is unfortunate that Fintoft has chosen to present the test words at the end of a carrier sentence. The frame jag såg ... där was used by Gårding and Lindblad (1973), where the speaker was asked to pronounce the sentences as neutral pieces of information. The same procedure was used in this test. All the words are recorded with the intonation pattern of a neutral statement<sup>3</sup>.

### The speakers

Speakers were selected according to the typology in Table 1. They were all considered to be typical representatives of their respective dialect.

#### Information about the speakers:

1A (Skåne): M, male, age 33. Grew up in Malmö. Parents from Skåne.

1B (Gotland): W, male, age 38. Grew up in Visby. Parents from Gotland.

2A (Stockholm): B, male, age 27. Grew up in Stockholm. Parents from Stockholm.

2B (Västergötland): L, male, age 39. Grew up in Västergötland. Parents from Västergötland.

Before the recording sessions the speakers read the sentences aloud and the meaning of each test word was discussed. The recordings were made in the language laboratory (learning studio) at the high school in Visby (Säveskolan) with the following recording equipment:

Beocord tape recorder, type 4132

Microphone AGK D 190 C

Agfa-Geavaert tape PE 21 LP (tape speed 19 cm/s)

The speakers listened to their speech afterwards and checked the recordings.

### The listeners

Listeners were taken from the high school in Visby. The selection was made at random and was dependent on the availability of pupils at the time the studio was free. Fintoft (1970, p. 53) also used pupils at secondary schools and says that since linguistic habits "are more or less fixed after the age of about 15, it is the region where the subjects lived during these first years of their life that matters when the regional homogeneity of the population is under consideration".

Number of listeners	Sex	Age	
4	male		
11	female	16	
16	male		
17	female	17	Total number
2	male		
15	female	18	of listeners
2	female	19	
1	female	20	68

Table 2. Distribution of listeners according to age and sex.

Table 2 shows the distribution of the listeners according to age and sex.

#### The listening test

The method of presenting the stimuli to the listeners has varied in other listening tests concerning isolated accent 1 and accent 2 words. Hadding (1961) used pictures for the identification. In Johansson's perceptual experiments (1970) the listeners (beginning students of phonetics) were instructed to mark each stimulus word "1" or "2". Fintoft (1970) used minimal pairs which were orthographically different. The listener's task was to identify the words as belonging to one of two specified phrases. Fintoft mentions that there may be some problems in identifying the phonemes when hearing short segments. In the present test both text and pictures were used, so that the listeners knew which segments they would hear before the actual identification (Fig. 2).

After the test words had been isolated from the frame sentence the material was transferred to another tape (two Beocord type 4132 tape recorders were used, tape speed 19 cm/s). The six test words appeared on the listening tape three times per speaker and were

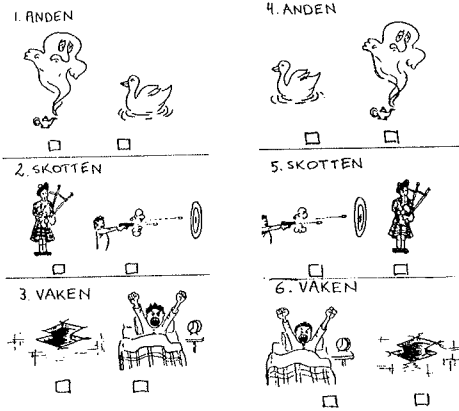


Fig. 2. Example of answer sheet.

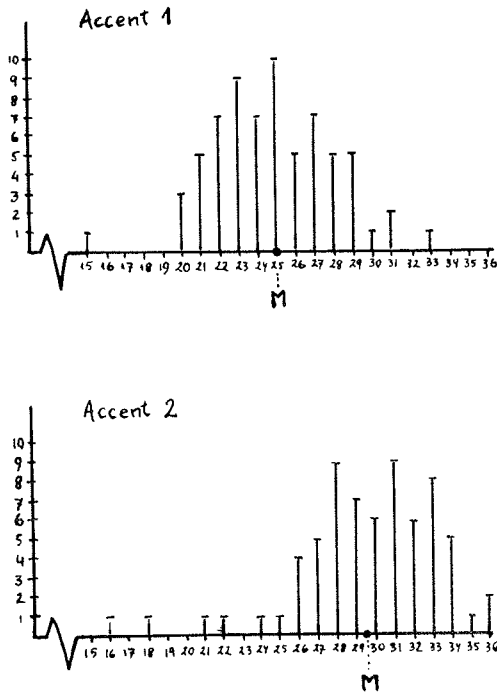


Fig. 3. Number of listeners (on the y axis) per correctly identified stimuli (on the x axis) and mean values (M) for accent 1 and 2.

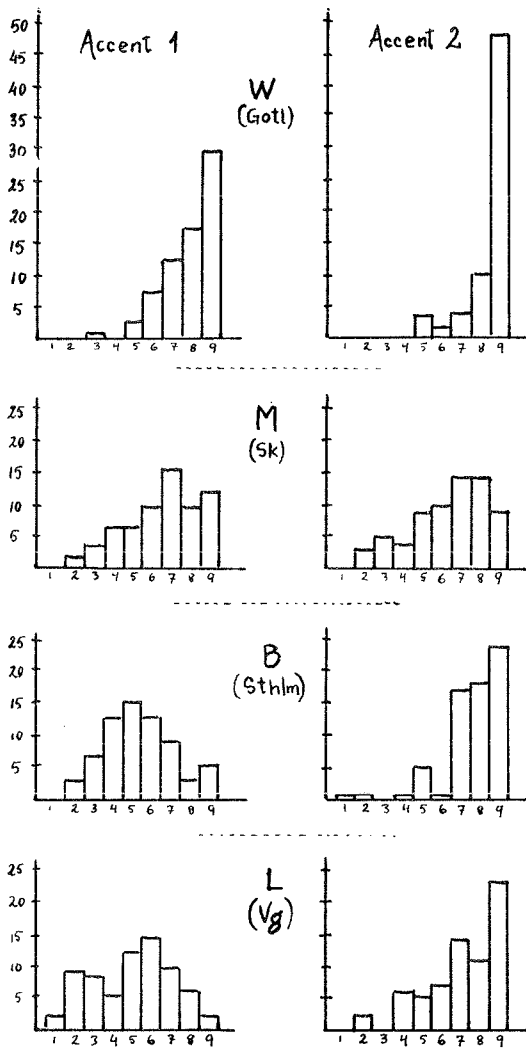


Fig. 4. Number of correct responses (on the y axis) per stimuli (on the x axis) for the accents of the speakers.



presented to the listeners in random order. Between words there was an interval of about 5 seconds and just before each word a ring was heard. After every tenth word a longer pause was made (10 seconds).

Before the test started the listeners were asked for personal data. They also wrote where they lived now and where they and their parents had grown up. In addition they read the following instruction: "You will hear a number of words. Before every word you will hear a ring. Give the meaning of each word by marking a cross in the appropriate square under each picture".

The listening test - with four testing sessions - took place in December 1974 in the language laboratory at Säveskolan, Visby. The tape recorder was the same as the one used when the material was recorded. Earphones: Primo Dynamic Head Phone, DH-843 L. After the test, which required nearly 15 minutes to complete, the listeners were asked to name the dialects they had recognized.

## RESULTS

77 subjects took part in the listening test. 68 of these had grown up on Gotland and with few exceptions had parents from the island. The other 9 had grown up in other parts of Sweden and their results were left out. This gave 4896 judgments, 72 per listener. The total number of correct responses was 3700, 1688 (68.9 %) for accent 1 and 2012 (82.1 %) for accent 2 (see histogram in Fig. 3). It has been mentioned that Hadding (1961) and Johansson (1970) reported higher scores for the accent 2 words. The same observation can be made from the present investigation. Table 3 and Fig. 4 show the distribution of correct responses for the different speakers.

In Hadding's (1961) investigation the identification scores depended on the listener's dialect and on his familiarity with the dialect of the test words. Johansson (1970) also found that the identification of isolated accents was dependent on the listener's own dialect pattern. Fintoft (1970) reported that listeners identified the word accents of their own dialects significantly correctly. The same can

	1A M (Sk)			1B W (Gotl)		
accent	1	2	1+2	1	2	1+2
correct responses	446	433	879	536	572	1108
%	72.9	70.8	71.8	87.6	93.5	90.5

	2A B (Sthlm)			2B L (Vg)		
accent	1	2	1+2	1	2	1+2
correct responses	362	517	879	344	490	834
%	59.2	84.5	71.8	56.2	80.1	68.0

Table 3. Number of correct responses per accent and speaker

be said of the present investigation. Speaker W (Gotland, 1B) received the highest recognition score, 90.5%. It should also be noted that 6 listeners were responsible 30% of the incorrect responses for W.

The figures for M (Skåne, 1A) show that M's accent 2 words did not get as high recognition scores as those of the other speakers. The relation between the scores for M's accent 1 and accent 2 words is more even than for the other speakers. As we have seen the pitch curves of Skåne and Gotland both have one peak in the two accent types. If the fundamental frequency pattern of one's own dialect is perceptually important this resemblance has obviously both helped and confused the listeners. The relatively high value for M's accent 1 indicates some degree of certainty. On the other hand the resemblance between the accent 2 pitch curve of Skåne and the accent 1 pitch curve of Gotland has made the listeners hesitate about M's accent 2 words compared to those of B and L.

Other factors are also perceptually important, at least concerning the present material. This is evident from the number of correct responses per word and from the scores for the accent 1 words of B (2A)-Stockholm. There is a certain resemblance between the accent 1 pitch curve of type 1B (shared by the listeners) and that of type 2A

(Stockholm), and one would expect high scores, which is also the case for the word vaken (accent 1 - see below). Furthermore, the pitch contour of accent 1, type 2B, resembles that of accent 2, type 1B (the listeners' own). The scores for L's accent 1 words are consequently low (56.2%) but the listeners have not consistently interpreted stimuli as accent 2. The fact that the fundamental is an important cue for recognizing the accents does not imply that it is the only factor involved in interdialectal perception. Malmberg (1959) points at the functional similarity and mentions that a native Swedish listener needs only hear a few sentences in an unfamiliar dialect before he can identify the accents. Fintoft (1970, p. 70) assumes that "listeners use different reference patterns for each dialect, which implies that listeners do not match the fundamental frequency patterns in other dialects with the patterns in their own. The assumption further implies that listeners estimate the dialect and match the fundamental frequency in a stimulus with the corresponding reference patterns. Sound quality differences between the dialects, and probably the frequency patterns in the complete words, may contribute to identify the dialects. Tonemes in dialects which are little known or unknown to the listeners seem to be interpreted according to the fundamental frequency patterns in the listeners' dialects. Responses to synthetic speech confirm this assumption".

In Fintoft's investigation many listeners obtained higher scores at the end of the test than at the beginning. In the present investigation 55 of the 68 listeners said they recognized dialect type 1B (Gotland - the listeners' own dialect), 48 type 1A (Skåne), 19 type 2A (Stockholm) and 5 type 2B (Västergötland)<sup>4</sup>. However, these groups have not made higher average identification scores than the others for the recognized dialects. According to Fintoft, most misinterpreted stimuli would occur at the beginning of the test, before the dialect had been recognized. But the mistakes are evenly spread over the whole test for all listeners.

The scores for accent 2 of each speaker are high compared to those for accent 1 (the only exception is M - 1A): W (1B) 93.5%, B (2A)

84.5%, L (2B) 80.1%, M (1A) 70.8% . Hadding (1961, p.71) tested a tentative assumption that the accent 2 of different Swedish dialects has some characteristics in common and found that "most unanimous correct responses were noted for Accent-2 words in all dialects". In a separate report of the test, Hadding (1962, p.637) writes that she found "an evident similarity between the tonal patterns of the stressed syllable of the Accent-2 words of the five dialects, viz., a distinct high, while the tonal patterns of Accent 1 were much vaguer and seemed to have no stable characteristics, except in southern Swedish".

Johansson (1970, p. 54) also noticed "a tendency to higher scores for the grave words. ... The grave accent is, however, generally considered to be marked, and in a number of perceptual experiments I have noticed that listeners tend to favour the marked member of an opposition, when a stimulus is felt to be somewhat 'strange'".

Malmberg (1962, p. 473) pointed out the relationship between the levels: "après avoir constaté le manque de correspondance entre la courbe de fréquence enregistrée et la forme mélodique perçue je me suis demandé si, malgré tout, il ne serait pas possible de trouver un point commun entre, disons, l'accent à Stockholm et le même accent dans le sud. ... Dans les deux dialectes, l'accent 1 est le type neutre (plus simple, moins spécifique), l'accent 2 le type caractérisé - le terme positif de l'opposition. C'est l'accent 2 qui est marqué ("merkmalhaftig"), l'accent 1 qui est non-marqué ("merkmallos"), ce qui est en parfait accord avec les faits fonctionnels".

Fintoft (1970, p. 276), discussed the concept of complexity and said that "it seems possible to regard the two tonemes as simple as opposed to complex. ... We may assume that the judgments of simple versus complex may differ somewhat from one dialect to another. The different potential cues may be more or less important to the different listener groups. We may also assume that one cue may override the effect of another. The judgement of simple versus complex is thus presumably based on the combination of all the

potential cues ... it seems as if there is some preference for toneme 2 when the frequency pattern may be characterized as complex".<sup>5</sup>

In the present investigation the listeners' identification score for accent 2 seems to be related to this complexity. The unknown dialects with an accent 2 pattern with two peaks have received the largest number of correct responses. All the listeners can be said to have experience of the dialect, type 2A (Stockholm), and to know it well enough from TV and radio speakers, teachers, visiting lecturers, friends who have moved from the region, travelling etc. Familiarity with the dialect cannot alone count for the high scores. In that case the number of correct responses for accent 1, type 2A, would have been larger. Instead, the listeners seem to identify the pitch pattern of each word by comparing it simultaneously to their own pattern and to their experience and knowledge of the dialect in question. The complex curves of accent 2 (two peaks) differ very much from accent 1 and from those of the listeners' group and are easily recognized. These remarks can also be applied to the result for L (2B).

The scores for M (1A) can be interpreted on this basis, too. There is a similarity between the fundamental frequency patterns of type 1A and type 1B (the listeners') and the scores suggest hesitation on the listeners' part. They have not known the pitch pattern of dialect 1A well enough (segmental differences may have helped them to recognize the dialect) and isolated words have been identified according to the listeners' own fundamental frequency pattern<sup>6</sup>.

No perceptually related difference was found between male and female listeners. Nor was there any difference between the results of the 28 listeners from Visby compared to those of the 40 listeners from other parts of the island.

Table 4 gives the identification scores in per cent for each word and speaker.

The scores for M (1A) show a strikingly small difference between the correct responses of the accents within each minimal pair (in no case over 2%). Another conspicuous feature is the relatively high values

	M(1A)	W(1B)	B(2A)	L(2B)
˘vaken 1	60.7	73.5	85.7	71.0
˘vaken 2	59.8	90.1	78.4	78.4
˘skotten 1	89.2	97.5	49.0	54.4
˘skotten 2	90.6	95.0	84.3	88.7
˘anden 1	66.6	91.1	42.1	41.6
˘anden 2	64.7	96.0	91.6	72.5

Table 4. Number of correct responses per stimulus and speaker (%)

of ˘skotten (1) and ˘skotten (2) for speaker M and for speaker W (the listeners' own dialect). The discussion above about the relationship between the listeners and their ability of accent identification is best illustrated with the scores of skotten (1 and 2) for all speakers: Accent 2 of B (2A) and L (2B) - complexity - obtains high scores (84.3 and 88.7%, respectively), whereas their accent 1 gets low scores (49 and 54.4%, respectively). The scores of anden (1 and 2) for B and L are similar, though the score for L's accent 2 is considerably lower (72.5). The listeners have, on the other hand, reacted differently to vaken (1 and 2). The scores for accent 1 are high (B 85.7% and L 71.0%). A reasonable explanation could be that the pitch curve of the long vowel was more easily perceived and related to a pattern (B (2A) - Stockholm) already known to the listeners.<sup>7</sup>

The figures of Table 4 may also indicate a connection between the degree of intelligibility of a word and its segmental phonetic manifestation. Fintoft (1970, p. 312) mentions this question: "It is reasonable to assume that potential cues such as duration and intensity variations and vowel quality, in some cases, may point in a direction other than the frequency cue, as these cues seem to be less stable than the  $f_0$  patterns. Particularly when listening to another dialect, it seems reasonable that some of these cues may override the  $f_0$  cue. The different features, not only the  $f_0$  patterns, seem to be different in the various dialects. Each dialect has apparently different rules for the assignment of the toneme pattern."



Fig. 5. Spectrograms of the test words vaken and anden (both accents).

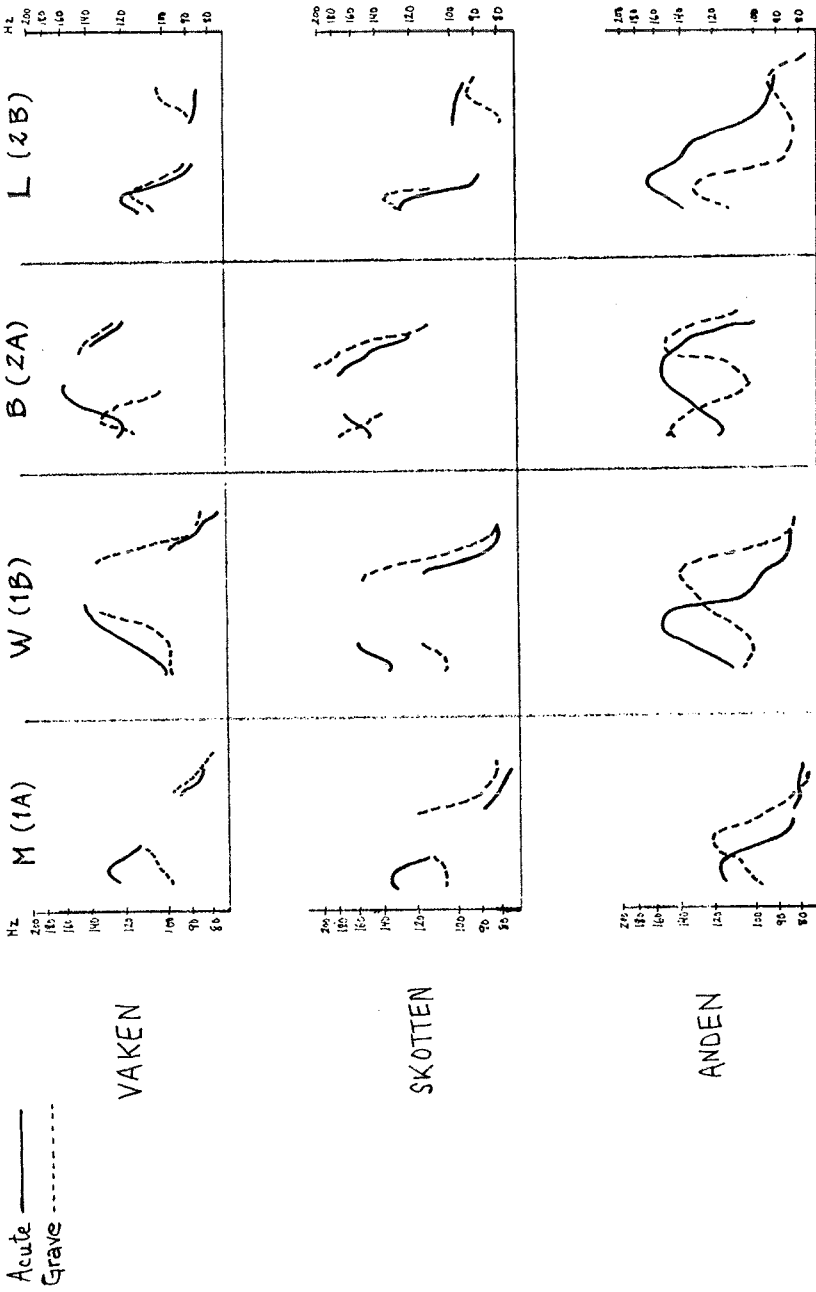


Fig. 6. Pitch curves of the test words traced from mingograms.



## DISCUSSION

Pitch patterns of the test words

Narrow band spectrograms of the test words vaken and anden are shown in Fig. 5 and  $f_0$  curves of all the test words, traced from mingograms, in Fig. 6. The different fundamental frequency patterns corroborate the results obtained by earlier investigators. The two peaks of B's and L's accent 2 words are clearly visible. The grave words of M and W have only one peak. Vaken (1) of W is not representative of the dialect. The rise of the contour is too slow and smooth and many listeners misinterpreted this word, although it was spoken by W, representing the listeners' own dialect.

A comparison between the identification scores and the pitch curves raises several questions. Is there a connection between the degree of intelligibility of a word and its segmental phonetic manifestation? One is apt to think that the "short" vowel and the broken  $f_0$  curve in skotten would make the word more difficult to recognize than the others. Instead, skotten obtained high scores, especially for M and W. Are certain segmental combinations more easily interpreted than others and are they dialectally related?

We do not know how dialects are recognized. The listeners who recognized the dialects were not able to get higher identification scores than the others, which is rather surprising. Thus it seems as if the recognition of a dialect is based on the different sound qualities, which then would be a more important cue than the fundamental frequency of the accents.

We have seen that the accent 2 words of B and L (acoustical complexity) got high identification scores. The listeners have been able to learn the accent 2 patterns of these dialects but have failed to respond correctly to the accent 1 patterns. Maybe this is connected with the functional similarity of the accents in the dialects. The constructive ('connective') function of accent 2 is present in every utterance. Since the distribution rules are uniform it is possible for a listener to master the accent 2 pattern perceptually after he has heard a few sentences of a dialect unknown to him. Is it then possible for him

to identify isolated words as well?

Various types of listening tests would probably throw more light on these and other questions concerning speaker-listener relations.

#### NOTES

1. The following facts were taken into consideration:  
The /r/ sound: The dialect of Skåne is the only one in the group with a fricative uvular [ʁ]. The dialect of Skåne has no supradentals. The dialect of Skåne and Gotland both have diphthongization of long vowels. The difference was considered to be least in /a/.
2. For vaken (awake) the frame jag låg...där (I lay...there) was used.
3. As the intonation pattern varies in the four dialects it is possible that the final fall of  $f_0$  in the frame sentence in some dialects may influence the pitch curve of the test word. Perceptually this is probably of no importance. (G. Bruce, personal communication.)
4. Segmental differences between the words may also serve as cues to the identification of a dialect.
5. Witting (1972) doubted the validity of the description for dialect type 1.
6. The responses from the listeners who were excluded from the investigation confirm Hadding's observation (1961, p. 70) that "listeners who had lived in more than one part of Sweden showed, as a group, better results than the group that did not report any such previous experience of more than one dialect".
7. The relatively few correct responses to W's vaken (1) (the listeners' own dialect) was caused by a stimulus that was not representative of accent 1 of the dialect.

#### REFERENCES

- Elert C.-C. 1970. Ljud och ord i svenskan. Stockholm
- Fintoft K. 1970. Acoustical analysis and perception of tonemes in some Norwegian dialects. Oslo

- Gårding E. 1973. The Scandinavian Word Accents. Working Papers 8, Phonetics Laboratory, Lund University
- Gårding E. 1974. Kontrastiv prosodi. Lund
- Gårding E. and P. Lindblad 1973. Constancy and variation in Swedish word accent patterns. Working Papers 7, Phonetics Laboratory, Lund University, 36-110
- Hadding-Koch K. 1961. Acoustico-phonetic Studies in the Intonation of Southern Swedish. Lund
- Hadding-Koch K. 1962. Notes on the Swedish Word Tones. Proceedings of the Fourth International Congress of Phonetic Sciences, Helsinki 1961. The Hague, 630-638
- Jensen M.K. 1961. Tonemicity. Acta Univ. Bergensis, Hum. Serie No. 1
- Johansson K. 1970. Perceptual experiments with Swedish disyllabic accent 1 and accent 2 words. Working Papers 3, Phonetics Laboratory, Lund University, 47-74
- Klintberg M. 1884. Laumålets kvantitet och aksent. Stockholm
- Malmberg B. 1955. Observations on the Swedish Word Accent. Mimeo-graphed report from Haskins Laboratories
- Malmberg B. 1959. Bemerkungen zum schwedischen Wortakzent. Zeitschrift für Phonetik und allgemeine Sprachwissenschaft 12, 193-207
- Malmberg B. 1962. Analyse instrumentale et structurale des faits d'accent. Proceedings of the Fourth International Congress of Phonetic Sciences, Helsinki 1961. The Hague, 456-475
- Meyer E.A. 1937. Die Intonation im Schwedischen I. Studies Scand. Philol. No. 10, Univ. Stockholm
- Meyer E.A. 1954. Die Intonation im Schwedischen II. Studies Scand. Philol. No. 11, Univ. Stockholm
- Pettersson R. 1953. Undersökningar av intonationen i gotländskan. Unpublished work. (Phonetics Laboratory, Lund University)
- Witting C. 1972. Nytt om ordintonation, särskilt norsk. Svenska landsmål och svenskt folkliiv. 38-51

Sidney Wood

A RADIOGRAPHIC ANALYSIS OF  
CONSTRICTION LOCATIONS FOR VOWELS

SUMMARY

Vocal tract area functions estimated from 38 sets of X-rayed vowel articulations collected from the literature and from new X-ray motion films of English and Arabic speech reveal four constriction locations: along the hard palate, along the soft palate, in the upper pharynx and in the lower pharynx. Each location is appropriate for a definable class of vowel qualities, confirming the quantal nature of at least this aspect of vowel articulation. The acoustical, physiological and phonological implications are discussed. In a given phonotactic environment the precision of the constricting tongue manoeuvre was good. The only truly language specific difference was a preference for either the midpalatal or prepalatal location for palatal constrictions. The tongue muscles are found to be admirably situated for creating constrictions at the four locations.

1 INTRODUCTION

Since the second half of the 19th century, work on vowel articulation has largely been based on a model that prescribes for each vowel a unique tongue position in terms of height and fronting of the tongue arch. This model, initiated by Bell (1867), was rapidly adopted long before it could ever be subjected to experimental verification. It has never been validated, but it has been contradicted. For example, Meyer's (1910) plastopalatograms revealed unexpected tongue heights. The Bell model was conclusively discredited by Russel's massive X-ray study (1928) which failed to corroborate the predicted tongue arch positions. This was followed by several decades of advances in acoustical analysis and psycho-acoustical experimentation, which has led to proposals for purely auditory or integrated acoustical-auditory systems for describing vowels (Jakobson et al. 1952, Ladefoged et al.

1972, Lindau 1975, Lieberman 1976). During the same half century articulation was largely disavowed for vowel description. The confusing picture of vowel articulation obtained from X-ray studies and the theoretical possibility of producing the same sound in a variety of ways were seen as apparent proofs of the inconstancy of articulation and the multiplicity of compensatory adjustments. However, I suggest in this paper that the alleged inconsistencies are due to observation of the wrong articulatory variables (height and fronting) rather than to articulatory irregularity. There are important and relevant regularities in vowel articulation (Wood 1975a), one of which will be reported here: the number of locations where the vocal tract is constricted by the tongue.

The location of a major internal narrowing is a fundamental parameter of vocal tract configuration for vowels (Stevens and House 1955, Fant 1960). These authors, and Lindblom and Sundberg (1971), treated the length of the vocal tract as a continuum of constriction location for model explorations of vocal tract resonances. On the other hand, the articulatory model universally accepted from ancient India until the 19th century had divided vowels into [i-ε]-like palatals, [u-o]-like labiovelars and [ɑ-a]-like pharyngeals. These are precisely the three regions where Stevens (1972) found vowel spectra to be relatively insensitive to moderate displacements of constriction location. Examination of three-parameter model nomograms (such as Fant 1960, Fig. 1.4-11) discloses four locations with this property: at the hard palate and in the lower pharynx for spread-lip vowels and at the soft palate and in the upper pharynx for rounded vowels (Gunnilstam 1974). I have previously inferred the same four locations from the formant transition frequencies of West Greenlandic Eskimo vowels (Wood 1971): (i) along the hard palate for the [i-e]-like allophones of /i/ and the [ε]-like allophone of /a/, (ii) along the soft palate for the [u]-like allophone of /u/, (iii) in the upper pharynx for the [o, ɔ]-like allophones of /u/ and the [ɣ]-like allophone of /i/, and (iv) in the lower pharynx for the [ɑ, a, æ]-like allophones of /a/. All this varied evidence points to the use of a small number of discrete loca-

tions for the constriction. The analysis of X-rayed vowel articulations reported below confirms these four constriction locations without exception by 40 subjects in 13 languages.

This finding has important implications for the study of speech physiology and the evolution of the speech organs. If it is true that it is not acoustically relevant to utilize more than four constriction locations for vowels, and if speakers universally confine themselves to these four locations, then the tongue must have a far simpler task than has hitherto been assumed. We should expect to find manoeuvres directed towards each of the four constriction targets in the vocal tract and a suitable arrangement of tongue musculature for this purpose.

## 2 RADIOGRAPHIC MATERIAL

This investigation is based on 38 sets of mid-sagittal vocal tract tracings for 12 languages that have been reported in the literature during the past 75 years (Wood 1975a) and on two new X-ray motion films of Southern British English and Egyptian Arabic speech. Tracings from a third film (of West Greenlandic Eskimo speech) are also used to illustrate part of this report. These films were made at the angiocardiological unit of the Lund University Hospital. The subjects were limited to one reel of 35 mm film each, which provided 40 seconds at 75 frames/second. The X-rays were delivered in one 3 msec pulse per frame. The absorbed radiation dose was 60-200 mrad/reel.

The English subject read a randomized list of test sentences of the form ['pVti 'pVt ə 'pVti, pVtə], where different test vowels were substituted for [V] and where [V] indicates the focus (intonation nucleus) of the sentence. The sentences were read at two different rates, yielding a total of 8 renderings of each of 10 different vowels. The Egyptian Arabic subject's sentences had the form ['bVti 'bVti 'bVta 'bVti] and ['bVṭa 'bVṭi] (with "emphatic" ṭ). These were read once and yielded 4 renderings of each vowel. For both subjects, the intervening lingual consonant and the weak vowels ensured that all test

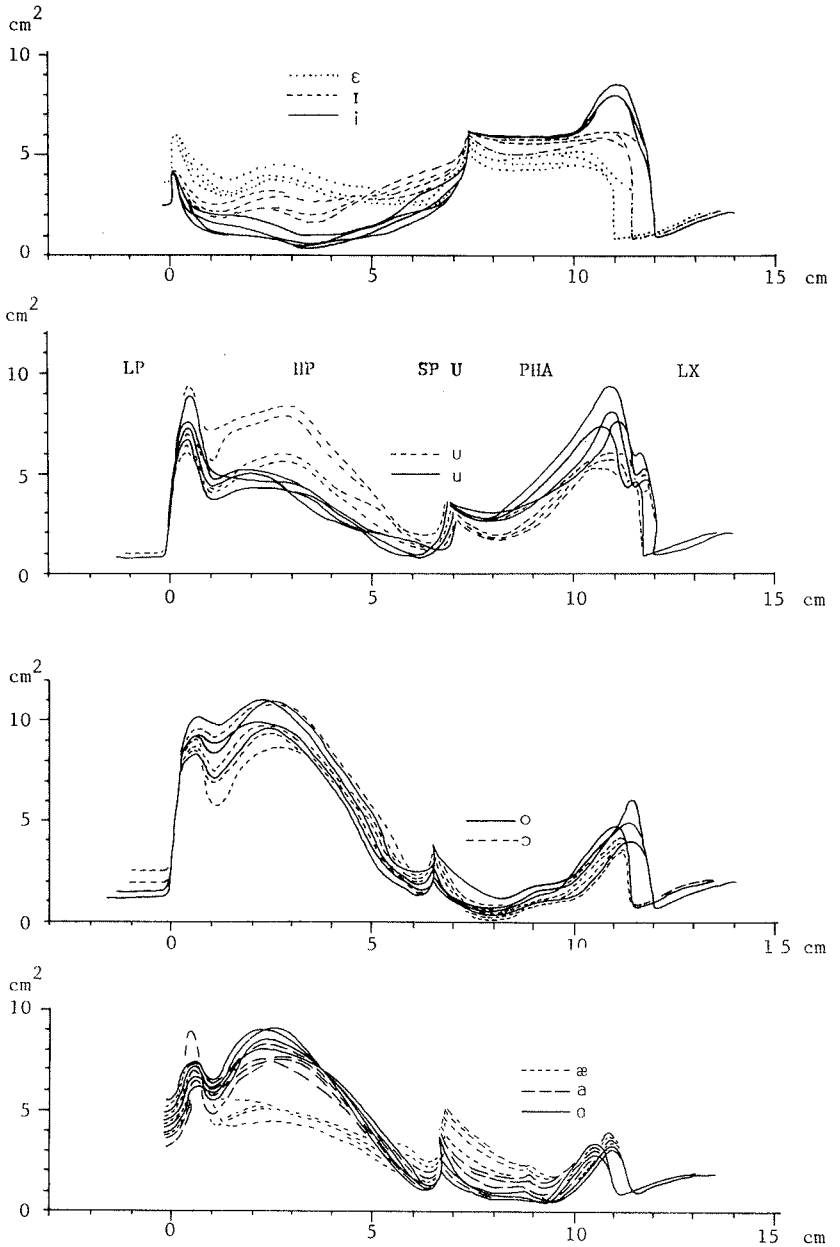


Fig. 1. The configuration of the vocal tract for vowels by the Southern British English subject, classed according to constriction location (hard palate, soft palate, upper pharynx, lower pharynx). The area functions are lined up at the central incisors (coordinate 0 cm). The letters identify parts of the vocal tract: LP lips, HP hard palate, SP soft palate, U uvula, PHA pharynx, LX larynx.

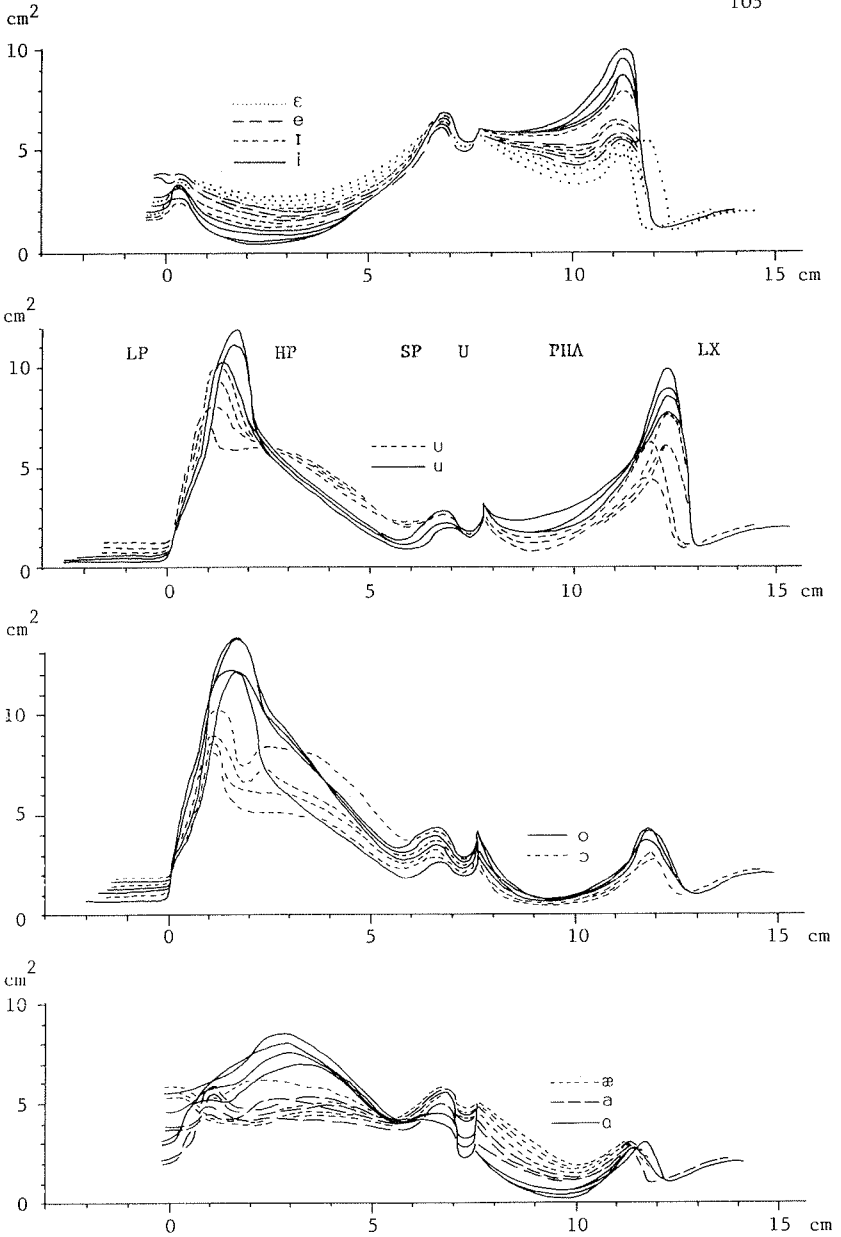


Fig. 2. Area functions for vowels by the Egyptian Arabic subject.



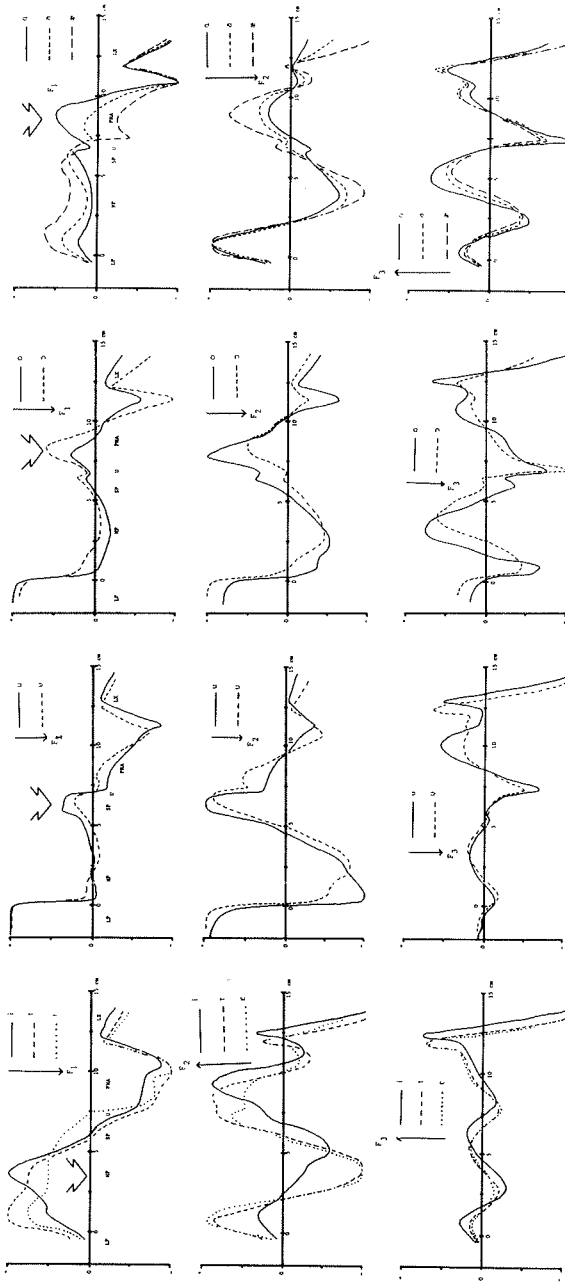


Fig. 3. The difference between the kinetic and potential energy distributions along the vocal tract, indicating the relative sensitivity of vocal tract resonances to local area perturbations. Local narrowing causes a resonance to fall or rise in proportion to the excess of kinetic or potential energy respectively (conversely for local expansion). English subject (cf. Fig. 1).

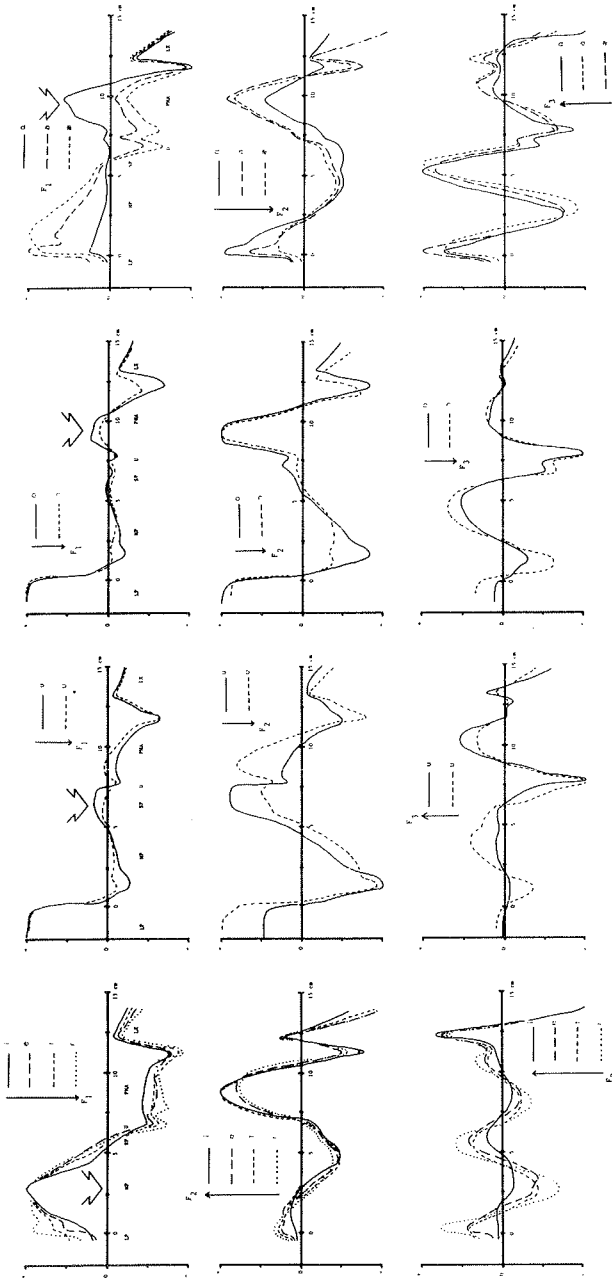


Fig. 4. Vocal tract resonance sensitivity functions, cf. Fig. 3. Egyptian Arabic subject (cf. Fig. 2).

that all test vowel articulations were independent. The number of tokens of each test item is important for statistical treatment (published sets of tracings rarely contain more than one token of each vowel).

The vocal tract area function for each rendering was estimated according to Sundberg (1969) and Lindblom and Sundberg (1971). Measurement of vocal resonance modes on a line electric analogue (LEA at the Speech Transmission Laboratory of the Stockholm Institute of Technology) indicated that estimated cross-section areas at the faucal isthmus in velar constrictions were too large. This will be discussed below (§ 6).

### 3 CONSTRICTION LOCATIONS

The striking tendency in all sets was that the tongue narrowed the vocal tract at one of four locations for vowels: (i) along the hard palate for [i-ε]-like and [y-ø]-like vowels, (ii) along the soft palate for [u-ʊ] and [ɨ]-like vowels, (iii) in the upper pharynx for [o-ɔ] and [ɣ]-like vowels, and (iv) in the lower pharynx for [ɑ-a-æ]-like vowels. Figs. 1 and 2 show area functions for my English and Arabic subjects. Area functions for the material collected from the literature were similar.

These are precisely those locations, referred to in the introduction, where resonance modes are insensitive to some displacement of the constriction. The fact that all subjects restricted themselves to these locations suggests that this preference is universal.

### 4 ACOUSTICAL CONSIDERATIONS

The combinations of  $F_1$  and  $F_2$  produced by constricting the vocal tract at each of the four locations can be studied by referring to three-parameter model nomograms. Fig. 5 is based on the Stevens and House (1955) nomograms and gives the frequencies of  $F_1$  and  $F_2$  generated by varying the degree of mouth-opening and the degree of constriction at each of the four constriction locations. Fig. 5 also contains the

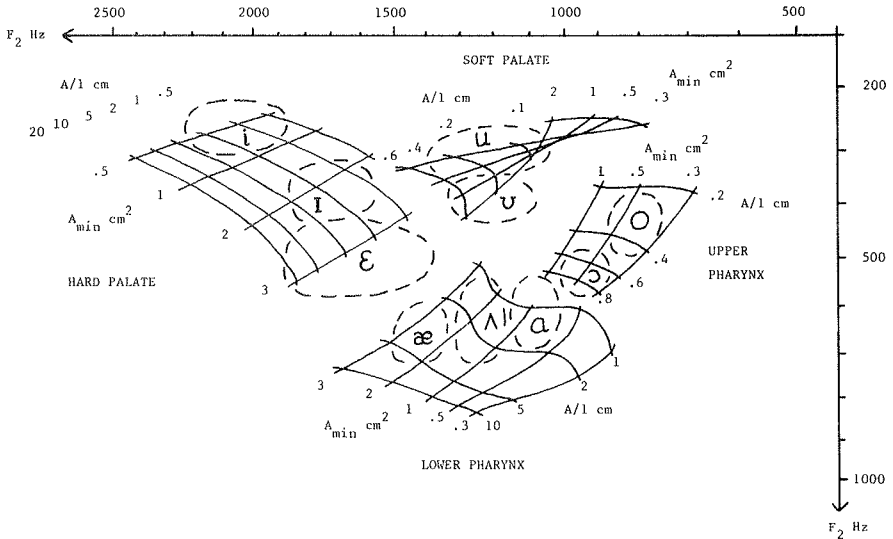


Fig. 5. The frequencies of  $F_1$  and  $F_2$  generated by the three-parameter model for the four preferred constriction locations, based on nomograms by Stevens and House (1955) (distance from source to constriction 12 cm for the hard palate, 8.5 cm for the soft palate, 6.5 cm for the upper pharynx, 4.5 cm for the lower pharynx). The superimposed vowel areas are from a sample of Southern British English speech recorded from the radio.

frequencies of  $F_1$  and  $F_2$  of stressed vowels in a sample of Southern British English speech recorded from the radio. The four constriction locations conveniently divide the entire  $F_1/F_2$  space into four relatively unambiguous areas, each enclosing a definable family of vowel qualities. Within each area, the different spectra are obtained by varying the degree of constriction (corresponding to lingual and mandibular articulation) and the degree of mouth-opening (corresponding to labial and mandibular articulation). In natural speech the formant frequencies are also determined by tongue root movement in the lower pharynx, by tongue blade movement in the buccal cavity and by vertical larynx movement. These movements cannot be simulated by a three-parameter model independently of the three model parameters. Independent simulation is only possible by direct manipulation of the area

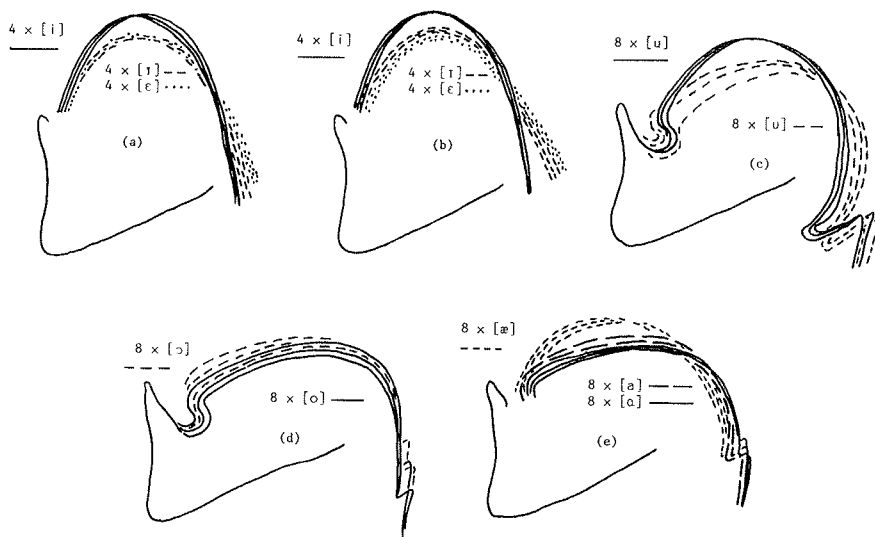


Fig. 6. Position of the tongue relative to the mandible for stressed vowels by the Southern British English subject, classed by constriction location. (a) slower rate, (b) faster rate, (c-e) both rates. The three profiles for each vowel represent the range of variation and the average position. Cf. Fig. 1. Note the mid-palatal posture for the palatal vowels (cf. Fig. 7).

function to reproduce the desired articulatory modification (Lindblom and Sundberg 1971, Mermelstein 1973).

The vocal tract is a single non-homogeneous pipe whose resonance modes are sensitive to local narrowing or expansion. All parts of the vocal tract contribute in varying degrees to each mode. The following laws help us understand the spectral consequences of each of the four preferred constriction locations. Firstly, local narrowing of the vocal tract will cause a resonance mode to rise or fall according as the perturbation is made in the vicinity of a sound pressure or volume velocity maximum in the standing wave for that mode (Chiba and Kajiyama 1941). The converse is true for local expansion. Secondly, the sensitivity of a resonance mode to a local area perturbation is related to the difference between the kinetic and potential energy at that point (Fant 1960, 1975; Schroeder 1967; Fant and Pauli 1975). A

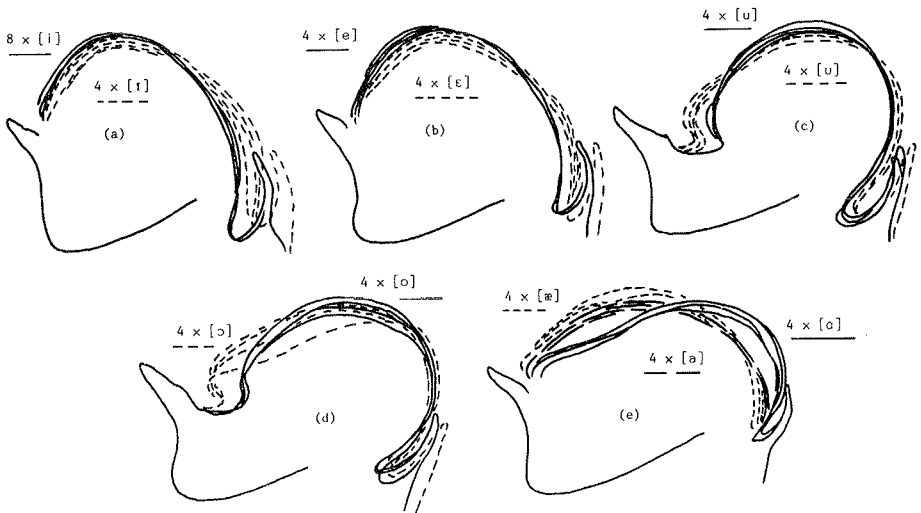


Fig. 7. The position of the tongue relative to the mandible for the stressed vowels by the Egyptian Arabic subject. Cf. Fig. 2. Note the prepalatal posture for the palatal vowels (cf. Fig. 6).

local expansion will cause a resonance mode to rise or fall in proportion to the excess of kinetic or potential energy respectively. Chiba and Kajiyama (1941), Fant (1960, 1975) and Mrayati and Carré (1976) give the volume velocity and sound pressure distributions for sets of Japanese, Russian and French vowels. In addition Fant (1975) and Mrayati and Carré (1976) give the energy distributions for their Russian and French sets. The sensitivity functions for my English and Arabic subjects are given at Figs. 3 and 4. All this data from different sources is strikingly similar for similar vowel qualities, as should be expected from the gross similarities of the corresponding areafuncions and the apparent universality of the four preferred constriction locations. It is therefore possible to generalize with confidence from this data to similar cases in other languages.

It is possible to predict from the volume velocity, sound pressure and

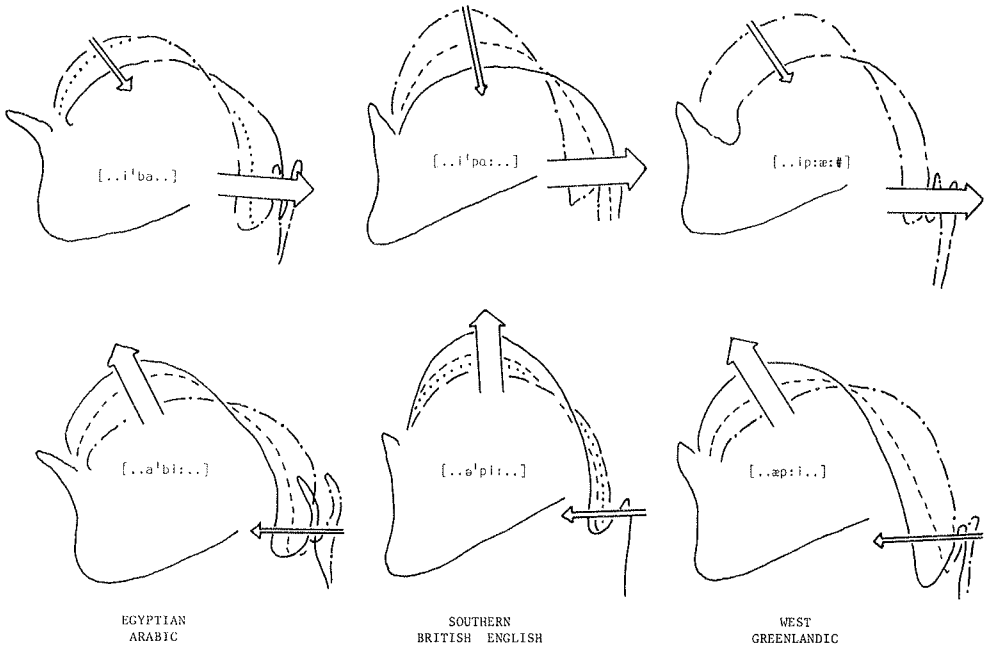


Fig. 8. The movement of the tongue relative to the mandible from a palatal [i], through [p, b] occlusion to low pharyngeal [a, ə, æ] (above) and vice-versa (below). The large arrows indicate the movement of the tongue body into the low pharyngeal or palatal constrictions respectively. The movement was sampled four times: first vowel — —, [p, b] occlusion ·····, [p, b] release - - - - , second vowel — —.

energy distributions that limited but specifiabale portions of the full F1 and F2 frequency ranges can be exploited at each of the four constriction locations (as exemplified at Fig. 5). These distributions also show why the formants are insensitive to slight constriction shifts at the four locations. The sound pressure and volume velocity maxima are not narrowly localized but range over extended zones. Consequently, resonance mode sensitivity to local narrowing or expansion does not alter appreciably through these zones (Figs. 3 and 4).

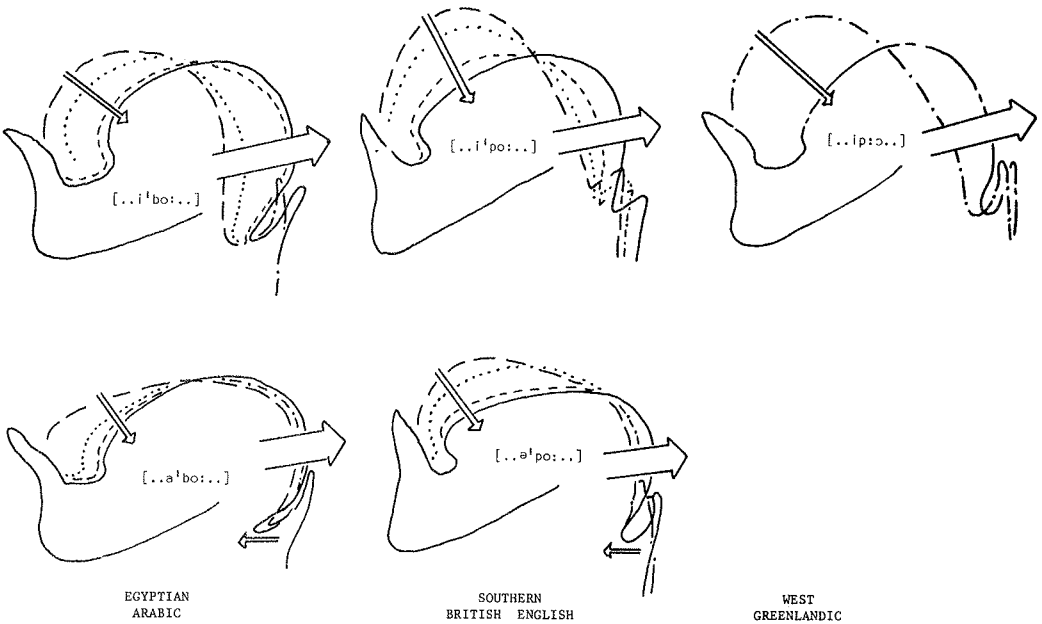


Fig. 9. The movement of the tongue relative to the mandible through [p, b] occlusion to upper pharyngeal [o] from a preceding [i] (above) and [a] (below). Cf. Fig. 8.

Since all parts of the vocal tract contribute to each resonance mode, vowel qualities within each of the four classes are not determined exclusively by varying the degree of constriction. It is particularly striking to note that in each of the four basic configuration types the tongue blade, the tongue body and the tongue root are in appropriate positions to narrow or expand the vocal tract precisely at sensitive parts of the F1 and F2 standing waves (Figs. 1, 2, 3, 4). Further, the extrinsic tongue muscles and the pharyngeal constrictors are admirably situated for this purpose (Fig. 11, § 7).



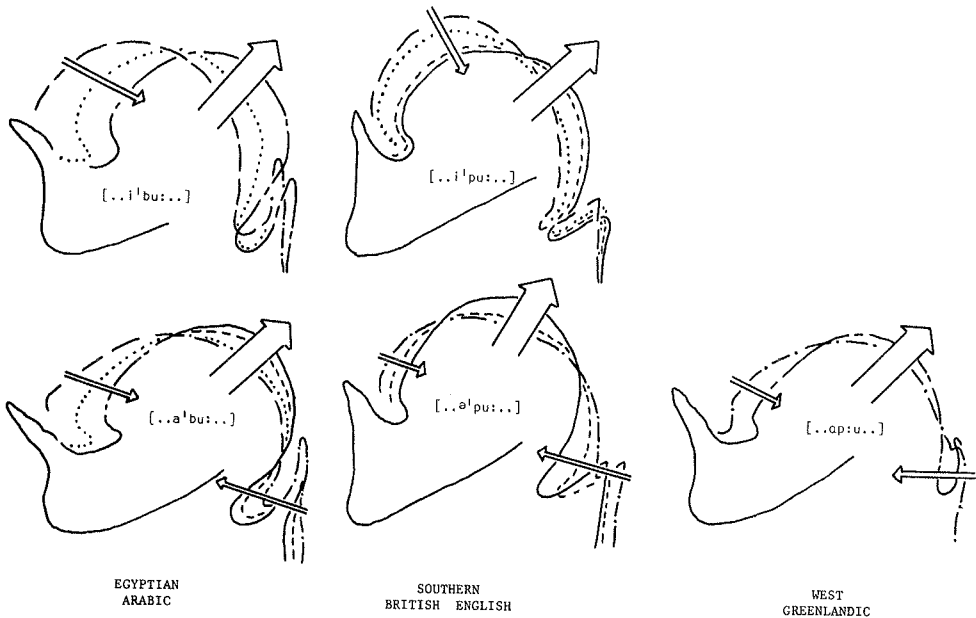


Fig. 10. The movement of the tongue relative to the mandible through [p, b] occlusion to velar [u] from a preceding palatal [i] (above) and low pharyngeal [a] (below). Cf. Fig. 8.

##### 5 THE PRECISION OF THE CONSTRICTING GESTURES AND THEIR RELATION TO LARYNX HEIGHT

Stevens (1972) has pointed out the possible articulatory advantage of utilizing the regions of the vocal tract where formant frequencies are insensitive to some variation of constriction location. Such variation might arise from the coarticulatory constraints of normal connected speech. We can imagine the speaker may strive to maintain the same distance from the glottal source to the constriction, implying lingual compensation for vertical larynx movement by fine adjustment of the direction of the constricting tongue gesture. Alternatively, the

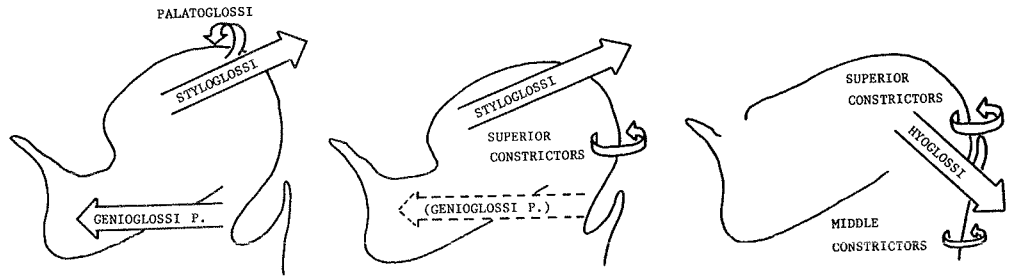
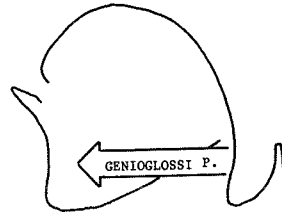
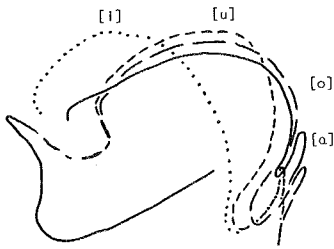


Fig. 11. The directions of contraction of the extrinsic muscles of the tongue and of the pharyngeal constrictors, arranged according to their presumed activity for the formation of the four constrictions.

speaker may constantly constrict the same part of the vocal tract, implying one location target disregarding the acoustical consequences of any variation of the vertical position of the larynx.

There is no evidence in the material reported here that speakers attempted to keep a constant distance from the glottal source to the constriction by making compensatory adjustments of the constricting gesture.

Figs. 1 and 2 are lined up on a fixed anatomical landmark, the central incisors. Variation of larynx height for any one vowel type is indicated by nonalignment at the glottal end. For these two subjects, the

tongue was directed towards the same part of the vocal tract for each token of a vowel type with little variation between tokens and irrespective of any variation of larynx height. A possible exception was the low pharyngeal constriction of the [ɑ, a, æ]-like class which tended to keep a constant distance to the glottis. This is probably because the upper part of the epiglottis contributes to this constriction (observed visually by Chiba and Kajiyama 1941 and Lindqvist and Sundberg 1971). The epiglottis is linked to the thyroid cartilage so that some of the vertical movement of the larynx is transmitted to it. Consequently, the constriction remains at about 4 or 5 cm from the glottal source.

Individual constricting tongue movements are illustrated at Figs. 8, 9, 10. These show the movement of the tongue from one vowel constriction to another. The similarities between the three subjects (representing three unrelated languages) should be noted.

It may be more pertinent to ask if there is any opportunity to vary the constriction locations. Some of the muscles involved have a sphincteral function with a localized constricting effect (the palatoglossi and the pharyngeal constrictors, see Fig. 11). Only in the case of the palatal constriction, where there are no muscles contracting across the vocal tract to pull the tongue upwards, does there seem to be any freedom for varying the location target. There is further discussion of this below (§§ 6, 7).

The range of larynx heights between separate tokens of each vowel type by both subjects was about  $\pm 2.5$  mm. The consequent variation of the distance from the glottal source to the constriction is within the acoustically tolerable limits and has little effect on the formant frequencies. The consequent variation of the overall length of the vocal tract also has little effect, about  $\pm 2\%$  for  $F_2$  of palatal vowels and much less for other vowels (Lindblom and Sundberg 1971, Wood 1975b). I assume that this random variation of larynx height from token to token reflects varying tensions of the extrinsic laryngeal musculature.

The larynx position does nevertheless appear to play an important roll in tuning the vocal tract. It is well known that the larynx is lower for rounded vowels than for spread lip vowels (Sundberg 1969). In the English and Arabic films the average larynx height was about 10 mm lower for "tense" [u, o] and about 5 mm lower for "lax" [u, o] than for spread-lip [i, e]-like vowels. In the tracings collected from the literature, the larynx tended to be lower for [u, o, y, ø]-like vowels than for [ɪ, γ, i, ε]-like vowels. By lengthening the vocal tract overall, this larger laryngeal depression contributes to the lowering of  $F_2$  of the rounded vowels, especially those with a palatal constriction. But more important is the adjustment of the distance from the glottal source to the constriction. Fant's three-parameter model nomograms clearly show that when the lip-opening is narrowed there is a lengthening of the distance from the glottal source to each of the four regions where formant frequencies are not sensitive to small shifts of constriction location. The invariant tongue constriction locations still coincide with these regions in rounded vowels thanks to the lower larynx. Statistically, these adjustments are quantal. The larynx is depressed for rounded vowels (more so for "tense" than for "lax" vowels) and elevated for spread-lip vowels.

#### 6 LANGUAGE-SPECIFIC DIFFERENCES OF CONSTRICTION LOCATION

The English and Arabic subjects had strikingly different constriction locations for palatal vowels (Figs. 1, 2, 6, 7). The English subject centred his constrictions midway along the hard palate about 35 mm behind the central incisors. The Arabic subject's constrictions were more anterior, about 27 mm behind the central incisors. This difference is reflected in the directions of their constricting tongue movements for palatal vowels (Figs. 6, 7, 8, 9). The American and British subjects in the collection of published tracings all had midpalatal constrictions centered at up to 40 mm from the incisors. Fant (1965) has pointed out a similar difference between "the Russian and Scandinavian [i] vowels which are prepalatal whereas the [i] is articulated more towards the midpalatal region in English", giving the

former a sharper quality. The consequence of the less anterior mid-palatal constriction is a wider prepalatal part and narrower postpalatal part, which will both yield a lower  $F_3$ . The  $F_3$  standing wave has a prepalatal pressure maximum and a postpalatal volume velocity maximum. Figs. 3 and 4 show that  $F_3$  is sensitive to modifications at both places in palatal vowels. It is interesting to note that most cases of prepalatal location in the collected material are from languages contrasting [i] with [y] or [ɨ] qualities. These contrasts are enhanced by having a maximally high  $F_3$  in [i].

The estimated area functions also indicated a difference between the velar constrictions for [u] by these two subjects (Figs. 1 and 2). The maximum narrowing estimated for the English subject was opposite the uvula, the vocal tract widening out anteriorly along the soft palate. The Arabic subject appeared to narrow the dorsovelar passage near the front end of the velum.  $F_2$  should not vary much for constriction shifts within this region but a model experiment on the analogue indicated that  $F_2$  would rise about 120 Hz when this constriction is fronted along the whole length of the velum. Yet paradoxically the English subject's  $F_2$  was the higher, 1000-1200 Hz against 750-900 Hz by the Arabic subject.  $F_2$  is usually higher than 1000 Hz for [u]-like vowels in English (cf. Fig. 5) and lower than 1000 Hz for the graver [u]-like vowels of some other languages such as Swedish. The English subject had less close lip-rounding and a less depressed tongue blade, which according to analogue simulation would together more than counter any negative effect on  $F_2$  from a slightly retracted constriction. But why is the Arabic subject's  $F_2$  so much lower? Other analogue experiments indicated that the constricted dorsovelar passage needs to be extremely narrow if  $F_2$  is to be lowered into the 700-1000 Hz range of a very grave [u] quality. Close lip-rounding alone is not sufficient. Narrowing of the dorsovelar passage is accompanied by narrowing of the top of the pharynx by the back of the tongue. The  $F_2$  standing wave has a volume velocity maximum extending right through the dorsovelar [u] constriction, with its peak in the pharynx immediately below the uvula.  $F_2$  is therefore lowered not only by

further narrowing of the dorsovelar passage but also by narrowing behind the constriction in the vicinity of the uvula.  $F_2$  of [u] has a considerable excess of kinetic energy throughout this region (Figs. 3 and 4) so that its frequency is extremely sensitive to narrowing all the way from the faucal isthmus to the top of the pharynx. A cross-section area in the constricted dorsovelar passage of at most 0.5 or 0.7  $\text{cm}^2$  would be necessary to keep  $F_2$  below 1000 Hz. The area given for the Arabic subject has probably been overestimated. It is notoriously difficult to assess cross-section areas accurately in the velar region from cross-distances alone. The palatoglossal muscles (in the anterior faucal pillars) can function sphincterally, which would narrow the faucal isthmus transversely and contribute to the degree of constriction (cf. Fig. 11). If these muscles are active for [u]-like vowels (the evidence will be discussed below in § 7), then they could help provide the extra narrowing needed for the Arabic subject to lower his  $F_2$  to well below 1000 Hz.

Of these two cases, the difference of location within the palatal region is most likely to provide a difference of vowel quality, by modifying  $F_3$ . For the velar region, any  $F_2$  variation attributable to differences of constriction location is much smaller than any opposite variation from other articulatory variables (degree of constriction, degree of lip-rounding, tongue blade depression, larynx depression) and is therefore easily cancelled by them. The true difference between the [u] vowels of these two subjects is the degree of constriction throughout the velar-uvular region.

The amount of freedom available for varying the constriction location is again a crucial problem in this context. The sphincteral mode of the palatoglossi will narrow the vocal tract locally in [u]. If the styloglossi and palatoglossi (Fig. 11) together guide the tongue into position for [u], little freedom should remain for varying the constriction location. For the palatal vowels on the other hand, there are no muscles that contract across the anterior part of the vocal tract to pull the tongue up into position. The tongue has to be pushed up from below, which leaves greater freedom for determining the target of the movement.

## 7 PHYSIOLOGICAL CONSIDERATIONS

This section is devoted to how the four constrictions may be achieved. The discussion is based entirely on published material - anatomical descriptions supplemented by reported EMG results. The conventional view, extending back at least to Hellwag (1781), is that the tongue body is positioned for vowels by its extrinsic muscles. Hellwag listed the genioglossi (for palatals), the styloglossi (for velars) and the hyoglossi (for pharyngeals). In addition, the vocal tract is shaped by the palatoglossi (linking the tongue to the velum), the superior pharyngeal constrictors (including the glossopharyngeal fibres) and the middle pharyngeal constrictors. There will also be a contribution from the intrinsic muscles, especially the inferior longitudinals for tongue blade depression. The directions of contraction ascribed to these muscles are illustrated at Fig. 11. EMG data is still limited and has largely become available during the past decade.

Figs. 6 and 7 show tracings of the tongue profile relative to the mandible for a selection of vowels at each location. This presentation highlights similarities and differences of purely lingual articulation. There was little variation between different tokens of the same vowel by the same speaker.

Figs. 8, 9, 10 show the movement of the tongue to the different constriction locations from a preceding weak [i], [ə] or [a] through [p] or [b] occlusion. These are examples of the lingual manoeuvres that form the constrictions. The manoeuvres by each speaker are strikingly similar although they are for three unrelated languages.

### Palatal vowels

There are no muscles that pull the tongue up towards the hard palate. The palatal tongue position is generally ascribed to contraction of the genioglossi, especially the posterior fibres. This is now being confirmed by EMG investigation (Harris 1971, Smith 1971, Raphael and Bell-Berti 1975, Miyawaki *et al.* 1975). These fibres pull the tongue root forward, widening the lower pharynx where both  $F_1$  and  $F_2$  are

sensitive (Figs. 3 and 4). This contraction is also believed to push the tongue body upwards towards the hard palate where it narrows the palatal passage. The formants are also sensitive to the degree of palatal constriction, especially  $F_1$ . The mylohyoid muscles will also elevate the tongue body by pushing it up from the floor of the mouth but there is little agreement among authors as to whether they are used for vowels. They may, for example, provide the extra lift for a consonantal obstruction against the roof of the mouth. Any additional help in shaping the tongue in this class must come from intrinsic lingual muscles.

#### Labiovelar vowels

The [u, ʊ]-like and [o, ɔ]-like vowels constitute two separate classes according to constriction location, but it is convenient to discuss them together here. Figs. 6, 7, 11 show how the tongue is lower and more retracted relative to the mandible for [o] than for [u]. The same difference is true for all subjects in the collection of published tracings. The wider jaw-opening for [o, ɔ] will help lower the tongue away from the soft palate but it cannot unaided transfer the constriction to the upper pharynx. This must be done by the tongue itself.

The styloglossi draw the tongue upwards and rearwards towards the uvula. EMG investigations by Harris (1971), Smith (1971), Raphael and Bell-Berti (1975) have found the styloglossi active for [u]-like vowels. The similar brunched-up tongue postures of [u] and [o] suggest the styloglossi are also active for [o]-like vowels, but the data is very limited. Smith found styloglossal activity in cardinal [o:] but hardly any in cardinal [ɔ:]. Miyawaki *et al.* (1975) also studied the styloglossi but their results were unfortunately marred by artefacts.

The styloglossi can determine the general upward and rearward direction of lingual movement. But how is the tongue elevated towards the soft palate for [u] or retracted into the pharynx for [o]? For elevation there are two possibilities - the genioglossi pushing from



below and the palatoglossi pulling from above. For retraction there are the superior pharyngeal constrictors. The hyoglossi, which draw the tongue down into the lower pharynx, have also been proposed in the literature for [o].

Harris (1971), Smith (1971), Raphael and Bell-Berti (1975) and Miyawaki *et al.* (1975) have all found the posterior genioglossal fibres to be active for [u]. Such activity can be deduced from the very wide lower pharynx for [u]. The posterior genioglossi should be less active for [o] since the tongue body is not elevated towards the velum and the tongue root is less advanced. At the same time there should be some activity for [o, ɔ] in these fibres in order to keep the lower pharynx sufficiently open for  $F_1$  to stay in the 350-550 Hz range. There is also a difference of tongue root position between "tense" [o] and "lax" [ɔ]-like vowels (Wood 1975b) which indicates a difference of genioglossal activity between them. The posterior fibres of the genioglossi have so far been studied for [o] by only one group of investigators, Miyawaki *et al.* (1975). They found that the same electrode that picked up considerable activity for [u] also detected activity for [o].

Let us now turn to the palatoglossi. These are a very slender pair of muscles linking the dorsal part of the tongue and the soft palate. They lie in the anterior faucal pillars. They appear to have several modes of activity depending on what other structures or muscles happen to be doing (Lubker 1975), including (i) to draw the tongue towards the soft palate when the latter is firmly elevated by the palatine levators, and (ii) to act like a sphincter to narrow the faucal isthmus transversely. Both of these modes are relevant for the articulation of [u]-like vowels by (i) guiding the velar (styloglossal) movement anteriorly towards the soft palate, and (ii) controlling the degree of constriction there. It is well known that the velum tends to be most tightly closed against the posterior pharyngeal wall during [u]-like vowels. An acoustic and perceptual explanation has usually been offered for this tight closure in an [u]-like configuration, based on the particular sensitivity of its oral resonances

to nasal excitation and on the fact that contrastive nasality is rare for [u]-like vowels. However, a velum tightly closed by the palatine levators is a necessary condition for the sphincteral mode of palatoglossal function, which means there is a firm physiological constraint against nasalizing [u]. Any EMG activity recorded from the palatoglossi during [u] can only be due to the two modes quoted above from Lubker. Other modes are not applicable for [u] (opening the oronasal passage when the palatine levators are not contracting, or adjusting muscle length to accommodate a downward movement of the tongue). Fritzell (1969) found strong bursts of potentials in the palatoglossi during the transition from [f] to [u] in foolish and in [u] in the sequences [bu, du, gu] and [mu, nu, nu] (all with strong simultaneous levator contraction). Of Bell-Berti and Hirose's (1973) two subjects, one revealed no palatoglossal activity whatsoever, neither for [u] nor for nasals. The other had palatoglossal activity for all nasal openings and [k] and the largest bursts of all for [u]. The balance of all this data is in favour of palatoglossal involvement in the articulation of [u] with simultaneous contraction of the palatine levators.

For the retraction of the tongue into the pharynx for [o, ɔ], the most likely muscles are the superior pharyngeal constrictors (including the glossopharyngeal fibres) in view of their upper pharyngeal situation. Smith (1971) found the glossopharyngei active in [o, ɔ]-like vowels. MacNeilage and Sholes (1964) concluded from the activity picked up by surface electrodes from the back of the tongue that the hyoglossi were involved for [o]. But surface activity at that location could equally well have come from the glossopharyngei. The hyoglossi have so far not been investigated with electrodes inserted into the muscles themselves and their involvement is open to speculation in the absence of further data. Two factors probably rule out the hyoglossi for [o, ɔ]: (i) the upper pharyngeal [o, ɔ] constriction occurs where the glossopharyngei and the superior pharyngeal constrictors embrace the pharynx, and (ii) the moderately large cavity in the lower pharynx for [o, ɔ], necessary to avoid an unduly high  $F_1$ , precludes hyoglossal activity.

I conclude that the tongue is most likely drawn upwards and rearwards by the styloglossi for both the [u]-like and the [o]-like classes, and that this common movement is deflected towards the soft palate by the palatoglossi and the posterior genioglossi for [u]-like vowels, and into the upper pharynx by the superior pharyngeal constrictors and the glossopharyngei for [o]-like vowels. The genioglossi are active to widen the lower pharynx and assist in raising the tongue, more so for [u]-like vowels than for [o]-like vowels. The sphincteral function of the palatoglossi and the pharyngeal constrictors probably leaves little freedom for variation of these constriction locations.

#### Low pharyngeal vowels

For the [ɔ, a, æ]-like vowels, there is considerable general narrowing of the lower pharynx where  $F_1$  has high pressure and  $F_2$  a sensitive volume velocity maximum. The maximum constriction is in the vicinity of the epiglottis. The acoustical consequences are that  $F_1$  is maximally high for all vowels in this class (at least 600-700 Hz with the large mouth-openings that are also typical of these vowels) and that the individual vowel qualities are distinguished by  $F_2$  according to the degree of constriction (cf. Fig. 5). In contrast, the lower pharynx is less narrow for [ɛ, ɔ] configurations for which  $F_1$  does not exceed about 550 Hz even with exaggerated mouth-openings.

The extrinsic tongue muscles that draw the tongue down into the lower pharynx are the hyoglossi. The inaccessibility of this pair has so far deterred investigators from attempting to insert electrodes. The pharyngeal constrictors are also frequently mentioned in conjunction with these vowels. The superior constrictors (including the glossopharyngeal fibres) retract the tongue and narrow the pharynx above the constriction. The middle constrictors, arising from the corni of the hyoid bones, narrow the pharynx at and below the constriction. Both Chiba and Kajiyama (1941) and Lindqvist and Sundberg (1971) have observed transversal narrowing of the constriction at the level of the epiglottis, indicating contraction of the constrictors. Smith (1971)

found considerable activity from the glossopharyngeal fibres for cardinal [bo:] and [ba:]. Minifie *et al.* (1974) recorded the activity of the superior and middle constrictors to investigate differences of "tenseness" and "laxness" in [p, b] and reported no apparent differences between the vowels [i] and [a]. The activity they observed during [a] is expected for constrictor involvement in this vowel, but that observed for [i] is unexpected for a palatal vowel. A previous ultrasonic scan (Minifie *et al.* 1970) had shown considerable inward displacement of the lateral pharyngeal walls (3-4 mm by the left wall) during the low pharyngeal vowels [ɑ, ʌ, æ] but little or nothing during [i, u], indicating constrictor activity for the former but not for the latter. Minifie *et al.* (1974) suggest that "had the EMG signals been integrated, discernable differences would have emerged" but they did not pursue the question further.

The availability of EMG data is least satisfactory of all in this class of vowels - nothing for the hyoglossi and occasionally contradictory for the pharyngeal constrictors. But the direct (visual) and indirect (X-ray, ultrasound) observations of movement are unanimous: the tongue is drawn into the lower pharynx and the pharynx is narrowed transversely by inward movement of the lateral walls. This points strongly to involvement of both the hyoglossi to depress and retract the tongue and of the pharyngeal constrictors to narrow the pharynx generally and especially at and below the constriction. Narrowing of the lower pharynx is predictable from acoustic theory to raise  $F_1$  to its maximum, above 600 or 700 Hz.

## 8 PHONOLOGICAL CONSIDERATIONS

All the data and discussion so far presented have clear implications for phonology. The preference for the four constriction locations is apparently universal. These locations are acoustically and physiologically significant. They divide the spectral space into four vowel quality families. The extrinsic tongue muscles are located just where they are needed to provide these constrictions and to narrow or expand the vocal tract at the sensitive parts of resonance mode standing

waves. Phonetic processes that claim to describe phonemic contrasts, allophonic distributions, vowel shifts, morphophonemic alternations and so on in articulatory terms should reflect this unique way of utilizing the innate acoustical and anatomical properties of the vocal tract. The established tongue articulation model, which professed to describe tongue articulation in terms of the height and fronting of the tongue arch, failed to do this.

Notice also that the boundaries between the four families (Fig. 5) constitute the basic phonemic contrasts. In two-phoneme systems such as Kabardian (Halle 1970) there is a contrast between (i) [A]-like vowels produced with low pharyngeal locations and (ii) non-[A]-like vowels produced at the other constriction locations. In three-phoneme systems there are contrasts between (i) [A]-like vowels with low pharyngeal constrictions, (ii) spread-lip [I-E]-like vowels with palatal constrictions, and (iii) rounded [U-O]-like vowels with velar and upper pharyngeal constrictions. Five-phoneme systems have contrasts between (i) [A]-like vowels with low-pharyngeal constrictions, (ii) spread-lip [I]-like and (iii) spread-lip [E]-like vowels with palatal constrictions, (iv) rounded [U]-like vowels with velar constrictions and (v) rounded [O]-like vowels with upper pharyngeal constrictions. No further contrasts are obtained by varying the constriction location. Other modifications are utilized for additional contrasts, for example the degree of constriction, the degree of mouth-opening and the degree of lip-rounding.

An interesting phonological problem is how far and under what conditions the phoneme boundaries just outlined may be transgressed. For example, in Arabic and Eskimo there are spread-lip [ɨ] or [ɣ]-like allophones of /i/ with velar or pharyngeal constrictions carried over from adjacent consonant articulations. Again, umlauting involves an interchange of constriction location between (i) low pharyngeal [a], upper pharyngeal [o] and velar [u] and (ii) palatal [e, ø, y] respectively. Similar interchanges of constriction locations are involved in instances of vowel harmony.

This approach to vowel articulation provides a new perspective to these types of phonological problem, whereas the established model cannot capture their true nature in terms of the height and fronting of the tongue arch. The ambiguities in the latter model regarding the relationships between parameter values, physiological activity, resonator configuration and spectral output means that it is powerless to explain central areas of speech production (Wood 1975a). The explanatory power of phonological theory will be greatly enhanced if the features of tongue articulation were instead derived from a comprehensive model based on the shaping of the entire vocal tract in a manner that directly relates physiology, articulation and acoustics.

## 9 CONCLUSIONS

1. There are four different places where the vocal tract is narrowly constricted by the tongue for vowels - along the hard palate, along the soft palate, in the upper pharynx and in the lower pharynx (Figs. 1, 2). This finding confirms Stevens's hypothesis that we seek to constrict the vocal tract for vowels at those places where  $F_1$  and  $F_2$  are least sensitive to variability of constriction location.
2. The vowels produced at these locations fall into distinct families: [i-ε, γ-ø]-like, [u-u, ṭ]-like, [o-ɔ, γ]-like and [ɑ-ə-æ]-like respectively (Figs. 1, 2, 3, 4, 5). This supports Stevens's conclusion regarding the quantal nature of vowels.
3. The tongue assumes characteristic postures relative to the mandible that correspond to the four constriction locations (Figs. 6, 7, 8, 9, 10). Lingual movements to these postures can be unambiguously referred to muscular activity (Fig. 11).
4. There are documented examples of languages preferring either the prepalatal or midpalatal locations for the palatal constriction. However, the sphincteral function of the palatoglossi and the pharyngeal constrictors leaves little opportunity to vary the locations of the other three constrictions.
5. In a given consonant environment there is good precision of the

constricting movements. There was no evidence that the direction of the constricting gesture is modified to compensate for random vertical larynx movement. But the structure of the epiglottis and the thyroid cartilage ensures that the low pharyngeal constriction automatically remains at about the same distance from the glottis.

6. When vowels are rounded, there is an increase in the distance from the glottis to each of the four regions where  $F_1$  and  $F_2$  are insensitive to small shifts of the constriction location. This is allowed for by depressing the larynx considerably for rounded vowels.

7. Articulatory features for use in phonology should reflect the preference for four constriction locations and the unique relationship between constricting tongue gestures, muscle situation and the degrees of sensitivity of vocal tract resonances to area perturbations at different parts of the vocal tract.

8. The approach to tongue articulation outlined here facilitates the building of a comprehensive description of speech production in which each of the successive stages (neuromotor, articulation, cavity, shaping, spectral output) are unambiguously related to each other. In contrast, the features of tongue arch height and fronting of the established model are ambiguous in these respects and constitute a capricious medium for relating the different phases of speech production.

#### Acknowledgments

I gratefully acknowledge the assistance of Gudmund Swahn, Gunnila Holje and Rolf Schoëner of the Roentgen Technology Section of the Lund University Hospital for the X-ray films, of Professor Gunnar Fant and Dr. Johan Sundberg of the Speech Transmission Laboratory at the Stockholm Institute of Technology for the use of the line analogue LEA, of Sven Norén of the Lund University Computer Centre for programming and of Dr. Eva Gårding and Dr. James Lubker for valuable discussion.

## REFERENCES

- Bell A.M. 1867. Visible speech. London
- Bell-Berti F. & H. Hirose. 1973. Patterns of palatoglossal activity and their implications for speech. SR 34:203-209. Haskins Laboratories
- Chiba T. & M. Kajiyama. 1941. The vowel - its nature and structure. Tokyo
- Fant C.G.M. 1960. The acoustic theory of speech production. The Hague: Mouton
- Fant C.G.M. 1965. Formants and cavities. Proceedings of the 5th International Congress of Phonetic Sciences, pp. 120-141. Basle: Karger
- Fant C.G.M. 1975. Vocal tract area and length perturbations. STL-QPSR 4/1975:1-14. Speech Transmission Laboratory, Institute of Technology, Stockholm
- Fant C.G.M. & S. Pauli. 1975. Spatial characteristics of vocal tract resonance modes. Speech communication, Vol. 2 pp. 121-132. Stockholm: Almqvist & Wiksell
- Fritzell B. 1969. The velopharyngeal muscles in speech. Acta Otolaryngologica, Suppl. 250
- Gunnilstam O. 1974. The theory of local linearity. Journal of Phonetics 2:91-108
- Halle M. 1970. Foundations of Language 6:95-103
- Harris K. 1971. Action of the extrinsic tongue musculature in the control of tongue position. SR 25/26:87-96, Haskins Laboratories
- Hellwag C.F. 1781. Dissertatio inauguralis physiologico-medica de formatione loquelae. Tübingen. Reprinted by Vietor 1886, Heilbronn
- Jakobson R., C.G.M. Fant & M. Halle. 1952. Preliminaries to speech analysis. Cambridge, Mass: MIT Press
- Ladefoged P., J.L. Declerk, M. Lindau & G. Papçun. 1972. An auditory-motor theory of speech production. UCLA Working Papers in Phonetics 22:48-75
- Lieberman P. 1976. Phonetic features and physiology: a reappraisal. Journal of Phonetics: 4:91-112



- Lindau M. 1975. Vowel Features. Working Papers 11:1-42. Department of Linguistics, Lund University
- Lindblom B.F. & J. Sundberg. 1971. Acoustical consequences of lip, tongue, jaw and larynx movements. *Journal of the Acoustical Society of America* 50:1166-1179
- Lindqvist J. & J. Sundberg. 1971. Pharyngeal constrictions. *STL-QPSR* 4/1971:26-31. Speech Transmission Laboratory, Institute of Technology, Stockholm
- Lubker J. 1975. Normal velopharyngeal function in speech. *Clinics in Plastic Surgery* 2:249-259
- MacNeilage P. & G.N. Sholes. 1964. An electromyographic study of the tongue during vowel production. *Journal of Speech and Hearing Research* 7:211-232
- Mermelstein P. 1973. Articulatory model for the study of speech production. *Journal of the Acoustical Society of America* 53: 1070-1082
- Meyer E. 1910. Untersuchungen über Lautbildung. *Festschrift Wilhelm Viëtor* pp. 166-248 (special number of *Die Neueren Sprachen*)
- Minifie F.D., J.H. Abbs, A. Tarlow & M. Kwaterski. 1974. EMG activity within the pharynx during speech production. *Journal of Speech and Hearing Research* 17:497-504
- Minifie F.D., T.J. Hixon, C.A. Kelsey & R.J. Woodhouse. 1970. Lateral pharyngeal wall movement during speech production. *Journal of Speech and Hearing Research* 13:584-594
- Miyawaki K., H. Hirose, T. Ushijima & M. Sawashima. 1975. A preliminary report on the electromyographic study of the activity of the lingual muscles. *Annual Bulletin of the Research Institute of Logopedics and Phonetics* 9:91-106. Tokyo
- Mrayati M. & R. Carré. 1976. Relations entre la forme du conduit vocal et les traits caractéristiques acoustiques des voyelles françaises. *Phonetica* 33:285-306
- Raphael L.J. & F. Bell-Berti. 1975. Tongue musculature and the feature of tension in English vowels. *Phonetica* 32:61-73
- Russel O. 1928. The vowel. Columbus, Ohio
- Schroeder M.R. 1967. Determination of the geometry of the human vocal tract by acoustic measurements. *Journal of the Acoustical Society of America* 41:1002-1010

- Smith T. 1971. A phonetic study of the function of the extrinsic tongue muscles. UCLA Working Papers in Phonetics 18
- Stevens K.N. 1972. The quantal nature of speech: evidence from articulatory-acoustic data. In David E.E. & P.B. Denes (eds.), Human Communication, a unified View pp. 51-66
- Stevens K.N. & A.S. House. 1955. Development of a quantitative description of vowel articulation. Journal of the Acoustical Society of America 27:484-495
- Sundberg J. 1969. Articulatory differences between sung and spoken vowels in singers. STL-QPSR 1/1969:33-45. Speech Transmission Laboratory, Institute of Technology, Stockholm
- Wood S. 1971. A spectrographic study of allophonic variation and vowel reduction in West Greenlandic Eskimo. Working Papers 4:58-94, Department of Linguistics, Lund University
- Wood S. 1975a. The weaknesses of the tongue-arching model of vowel articulation. Working Papers 11:55-108. Department of Linguistics, Lund University
- Wood S. 1975b. Tense and lax vowels - degree of constriction or pharyngeal volume? Working Papers 11:109-134. Department of Linguistics, Lund University



TEMPORAL ORGANIZATION OF SWEDISH TONAL ACCENTS: THE EFFECT OF VOWEL  
DURATION IN THE GOTLAND DIALECT

Robert Bannert and Anne-Christine Bredvad-Jensen

1 INTRODUCTION

The data presented here are an addition to the investigation of Bannert & Bredvad-Jensen (1975). In order to study the effect of vowel duration on the tonal accent contours of all the four tonal types of Swedish dialects proposed by Gårding (1973), recordings of a speaker from Gotland, the dialect not represented in the previous study, were made afterwards. This dialect represents type 1B of Gårding's typology.

The test material is the same as in the first study (cf. p. 3, all page references are to Bannert & Bredvad-Jensen, 1975), but we added four test words to achieve an optimal segment structure correspondence between words with a long stressed vowel compared to words with a short stressed vowel. The four words are:

<u>accent 1</u>	<u>accent 2</u>
<u>hack-en</u> (def. pl. of "notch")	<u>hacken</u> (solemn imperative of 2nd pers. pl. of hacka = "to chop")
<u>stick-en</u> (def. pl. of "stitch")	<u>sticken</u> (solemn imperative of 2nd pers. pl. of sticka = "to stick")

The recordings, with informant 1B, male, 21 years, and the acoustic analysis, were made in the same way as before (cf. p. 4).

2 HYPOTHESES AND PREVIOUS RESULTS

The investigation was designed to check whether the main  $F_0$ -changes would be attributed either to the truncation or to the compression hypotheses (cf. p. 1 f), when the stressed vowel in a bisyllabic word was radically shortened. The data from our seven informants supported both hypotheses (cf. p. 15). In most cases the short contours were derived from the long ones by the principle of truncation irrespective of the shape of the  $F_0$  contours of the word accents.

## 3 RESULTS AND DISCUSSION

The results are presented in the same way as before. Table 9 gives the  $F_0$ -values and the durations. Fig. 13 shows simplified  $F_0$ -contours of the first vowel segment in the twenty test words. For the arrangement see p. 7. Fig. 14a shows the  $F_0$ -tracings of tagen and tacken, both accents. In Fig. 14b these tracings are superimposed.

Durations. The stressed vowel segment of accent 1 is shorter with respect to its long counterpart than that of accent 2:

Accent 1			Accent 2		
Mean duration		V/V:	Mean duration		V/V:
ms			ms		
V:	V	%	V:	V	%
161	92	57	161	99	62

This informant behaves contrarily to the other seven informants (cf. Table 1, p. 8).

$F_0$ -contours. Manifestation in the long vowels: the  $F_0$ -curve of accent 1 is characterized by a rise to a  $F_0$ -maximum, which is followed by a fall during the last third of the vowel. The manifestation of accent 2 begins with a slight fall to a  $F_0$ -minimum in the first part of the vowel. It is followed by a rise throughout the rest of the vowel. The  $F_0$ -maximum is reached in the post accent consonant (Fig. 14a).

Vowel shortening: When vowel duration is decreased, the final fall of the long accent 1 contour is missing except in sticken and stucken, where the curves are compressed. Compression is found in all accent 2 vowels but one, skotten (Fig. 13).

Thus the effect of vowel shortening is not uniform for this informant, as was found to be the case with other informants as well. It seems as if in accent 1 truncation is mainly used whereas compression is used in accent 2 (cf. Table 9, p. 15).

Considering the data from our eight informants representing all the four tonal dialect types, it is surprising to find so little consistency in the tonal behaviour. But as we remarked finally in our previous paper (p. 17) the  $F_0$ -contours should be studied in a wider perspective. The prosodic components manifested in the  $F_0$ -contours should be controlled (cf. Bruce, forthcoming). Then an explanation for this seemingly contradictory behaviour may be found.

#### REFERENCES

- Bannert R. & A.-C. Bredvad-Jensen. 1975. Temporal organization of Swedish tonal accents: The effect of vowel duration. Working Papers 10, 1-36. Phonetics Laboratory and Department of General Linguistics, Lund University
- Bruce G. Forthcoming. Swedish word accents in sentence perspective. Travaux de l'Institut de Linguistique de Lund. B. Malmberg and K. Hadding (eds.)
- Gårding E. 1973. The Scandinavian word accents. Working Papers 8. Phonetics Laboratory, Lund University

Table 9. GOTLAND

F<sub>o</sub>-values (Hz) and duration t (ms)

		Accent 1				Accent 2			
tagen	F <sub>o</sub> :	-	121	134	124	-	109	99	119
	t :	0	18	122	174	0	24	97	192
tacken	F <sub>o</sub> :	-	117		137	-	112	107	124
	t :	0	22		98	0	16	58	108
hagen	F <sub>o</sub> :	123		132	121	105		99	118
	t :	0		119	165	0		59	156
hacken	F <sub>o</sub> :	118			137	113		106	120
	t :	0			83	0		42	93
backen	F <sub>o</sub> :	107			125	-	96	99	122
	t :	0			100	0	21	44	106
stigen	F <sub>o</sub> :	115		125	115	110		105	120
	t :	0		105	155	0		50	145
sticken	F <sub>o</sub> :	-	120	120	115	-	110	105	120
	t :	0	15	50	85	0	20	45	100
stucken	F <sub>o</sub> :	-	120	125	120	-	110	110	120
	t :	0	15	55	90	0	15	35	90
stegen	F <sub>o</sub> :	120		125	115	110		105	120
	t :	0		100	150	0		45	150
skotten	F <sub>o</sub> :	-	120	125	125	-	105		120
	t :	0	15	80	95	0	20		95

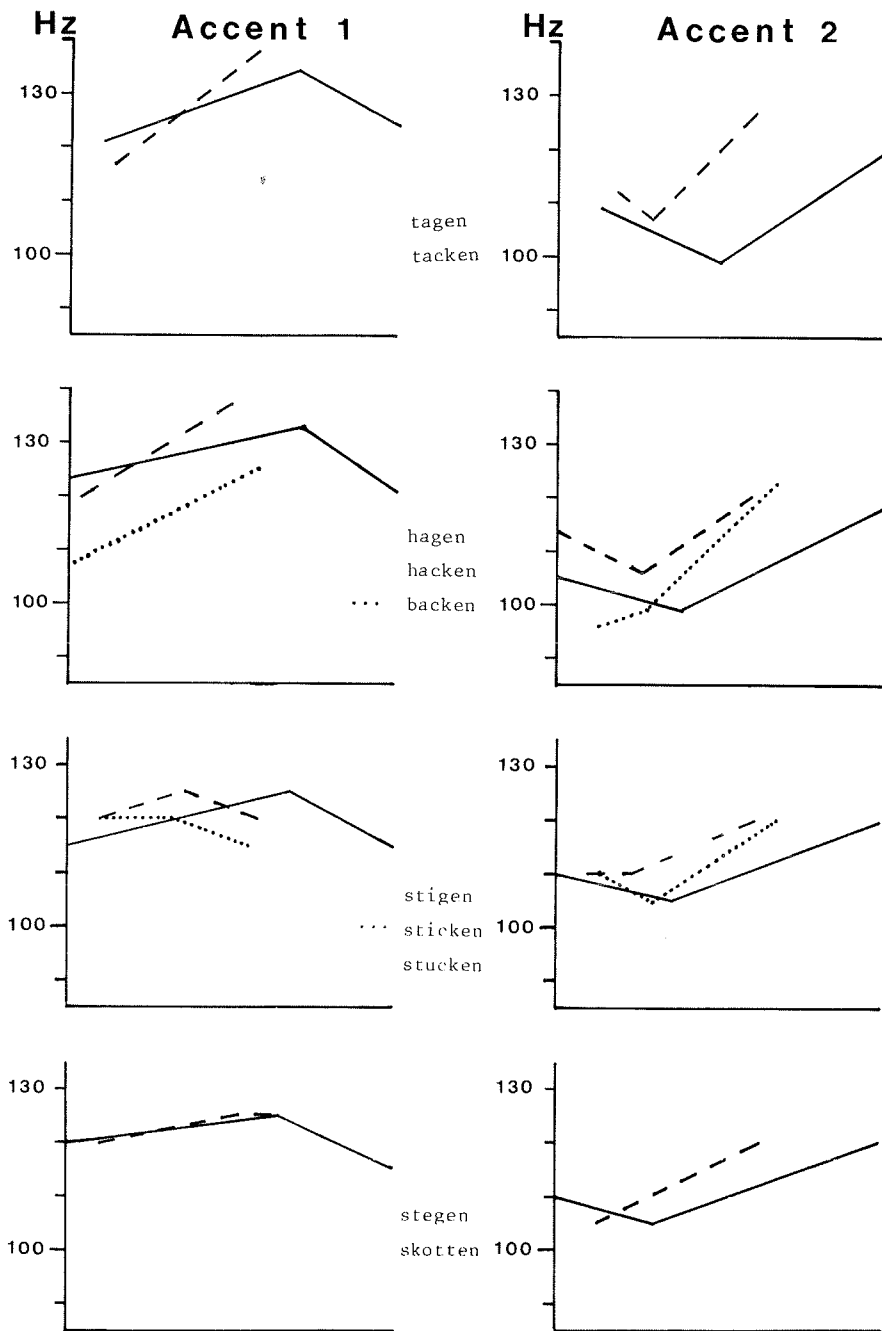


Figure 13. Simplified  $F_0$ -contours of the first vowel segment in the 20 test words. GOTLAND. The duration of the long vowels is normalized (solid lines). Dashed and dotted lines: short vowels.



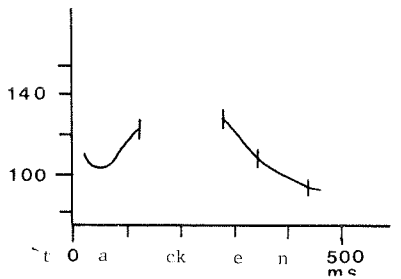
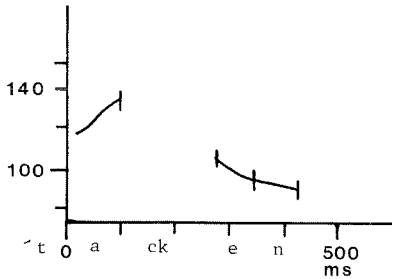
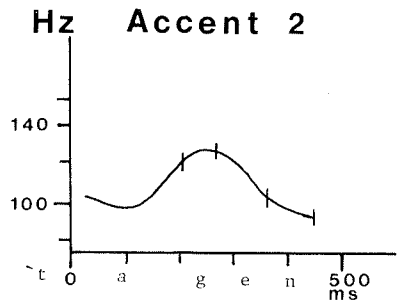
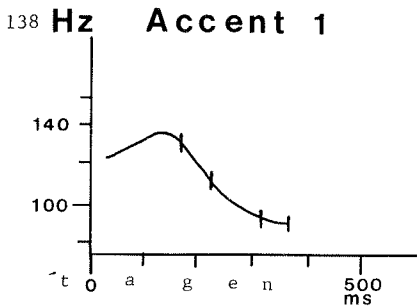


Figure 14a.  $F_0$ -tracings of tagen and tacken, both accents. GOTLAND

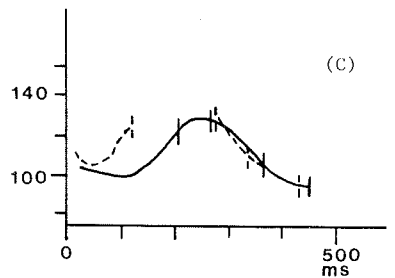
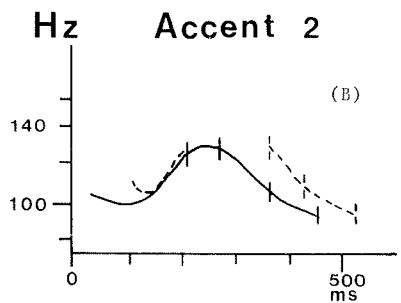
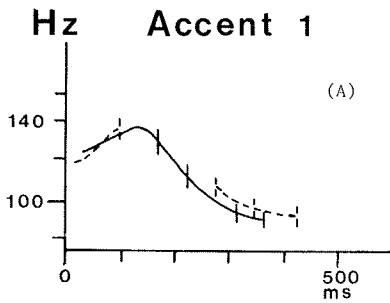


Figure 14b.  $F_0$ -tracings of tagen and tacken, both accents, short superimposed on long. Line-up point for A and C is the beginning of the vowel. For accent 2 (B) the curves are given so that the rise of the long and short stressed vowels coincide. GOTLAND.