FURTHER WORK ON COMPUTER TESTING OF A GENERATIVE PHONOLOGY OF DANISH

Hans Basbøll and Kjeld Kristensen[1]

1. Introduction

The purpose of the project to be reported here (called "DANFON") is that of a computational testing of a generative phonology of Danish (which, in its main lines, was already worked out by one of the authors, HB, before the project started). The ultimate purpose is that of improving the generative description of Danish phonology which is being tested, and also that of expanding the coverage of this phonology (see further section 5 below). The reader is referred to Basbøll and Kristensen 1974 for an account of the general structure of program and organization of data. Only points which were unmentioned in the previous report, or which have been changed since this report, will be included in the following (very preliminary) survey.

2. Some phonological aspects of the project

A string in abstract phonological representation is the input, and the program then changes it to an output (or several outputs) in phonetic (IPA) notation, by successive application of the phonological rules contained within the "grammar". The grammar makes use of three sets of background data, viz. UNIT-MATRIX (whose two dimensions consist of the distinctive features

---

1) Kjeld Kristensen is an engineer and cand. phil. in Danish. He teaches at the Institute of Scandinavian studies (University of Copenhagen). - We are indebted to Peter Holtse for valuable discussions, and to Jørgen Rischel for stylistic suggestions.

and the inventory of units, i.e. phonetic segments and boundaries, see sections 2 - 2.1 below), RULEMATRIX (whose two dimensions consist of the distinctive features and the (incompletely specified) units defining the structural description and change of each phonological rule), and RULEINDEX (which gives general information on each phonological rule concerning optionality, sensitivity to syllable boundaries, and the location of the rule in RULEMATRIX). See section 3.2 below for the description of a RUN, and, for further details, Basbøll and Kristensen 1974.

The phonetic notation used for the output forms must be narrow, since we want to be able to distinguish between all stylistically relevant phonetic differences (due to our interest in the possible hierarchy of optional rules). We therefore operate with 89 distinct phonetic segments ("sound symbols"), for the moment. These are listed below (in IPA-notation; for practical reasons[1] we are forced to use wrong symbols in some cases which will be indicated directly below). The first sixteen symbols (viz. those denoting full vowels) occur both as long and short vowels, and both with and without stød (see section 2.3 below); the number of non-composite IPA-symbols below therefore reduces to 89 - 3x16 = 41. The grammar operates with two further units, viz. the syllable boundary ($ on the line printer, ˅ on the IPA ball-head) and a grammatical boundary (# on the line printer, / on the IPA ball-head), see further section 2.1 below. Notice that all these phonetic segments and boundaries are used on most levels of the derivation: many of them occur as abstract phonological segments, and all of them as phonetic segments (and as intermediate segments).

---

1) Partly because the correct symbols are not found on the IPA ball-head, partly because certain characters on the papertape typewriter are prohibited in normal use (for computational reasons).

Phonetic segments:

(1) 16 full vowels:

i e ɛ æ ɑ

(e.g. in bil 'car', lidt 'little, adv.', fæ 'fool', hane 'cock', varm 'hot' [bi·?l, led, fɛ·?, hæ·nə, vɑ·?m])

y ø œ θ (for Œ, i.e. a low rounded front vowel)

(e.g. in dyst 'fight', løb 'run', høne 'hen', grøn 'green' [dysd, lø·?b, hœ·nə, gʁŒ n?])

u o ɔ ɒ (for ɒ)

(e.g. in mus 'mouse', kone 'wife', blå 'blue', år 'year' [mu·?s, ko·nə, blɔ·?, ɒ·?])

a ɤ (for ɑ, i.e. a vowel intermediate between a and ɑ) ʌ

(e.g. in land 'country', lam 'lamb', hånd 'hand' [lan?, lɑm?, hʌn?])

(2) 3 weak vowels:

ɪ[1] ɨ[2] ə

(e.g. in dydig 'virtuous', madding 'bait', hoppe 'hop' [dy·ðɪ, maðɨŋ, hʌbə])

---

1) The symbol ɪ denotes a "weak i" (derived from /ə/ by vowel raising before velars) which is found in the derivative ending -ig. If this sound is phonetically identical with a normal [i], a late tensing rule: ɪ → i may be included in the grammar (cf. the following footnote).

2) The symbol ɨ denotes a "weak e" (derived from ɪ by vowel lowering before nasals) which is found in the derivative ending -ing. If this sound is phonetically identical with a normal [e], a late tensing rule: ɨ → e may be included in the grammar (cf. the preceding footnote).

(3) 3 non-syllabic components of diphthongs:

j w ɹ

(e.g. in jeg 'I', tov 'rope', bær 'berry'
[j ∝j, tɒw, bæɹ])

(4) 13 obstruents:

p t k

(e.g. in på 'on', te 'tea', kom 'came'
pɔ·?, te·?, kʌm?])

f s ʃ h

(e.g. in få 'get', så 'saw', sjæl 'soul', hund 'dog'
[fɔ·?, sɔ·?, ʃɛ·?l, hun?])

b d g

(e.g. in ben 'bone', dyr 'animal', gå 'walk'
[be·?n, dy·?ɹ, gɔ·?])

v ʁ ʀ (for ʁ̥)

(e.g. in vild 'wild', rå 'raw', kors 'cross'
[vil?, ʁɔ·?, kɒʁ̥s[1]])

---

1) We have chosen to operate with a separate symbol for the unvoiced ʁ in the sequences rp, rt, rk, rf, rs (as a phonetic notation, this applies only in very conservative standards, of course). The point is that these sequences generally do not have "stød-basis" in conservative standards (this state of affairs is now in the process of change); compare the fact that words like kors [kɒ·s] in modern pronunciation violate the general restriction that monosyllables with a long vowel have stød. Notice, however, that we use the normal r-symbol, viz. ʁ, in the phonetic notation of words like pris 'price' [pʁi·?s] which phonetically have unvoiced ʁ, just as we use the normal (voiced) symbols [j l v], etc. also after [p t k f s] in words like pjat 'nonsense', klo 'claw', tværs 'across' [pjad, klo·?, tvæʁ̥s/tvæɹs]. This is because we consider (in agreement with Peter Holtse) the devoicing after "aspirates" to be a purely phonetic process, the aspiration phase being concurrent with the articulation of the following consonant.

(5) 6 non-syllabic sonorants:

l m n ɳ

(e.g. in lys 'light', mus 'mouse', nå 'reach', lang 'long' [ly·?s, mu·?s, nɔ·?, l ɑ ŋ?·])

ð ɣ

(e.g. in fed 'fat', fag 'profession' [fe·?ð, fæ·?ɣ])

2.1 Distinctive features

The 89 phonetic segments listed above, together with the two boundaries (and the blank), are cross-classified by 18 distinctive features which will be mentioned below. The choice of features as well as the feature analysis of the segments must be considered very preliminary. In particular, we may want to change this part of the grammar as a result of the attempted coordination with Peter Holtse's project in progress of synthesis by rule of Standard Danish. Reference to the voluminous literature on distinctive features will generally be omitted here.

The 18 distinctive features are the following:

Unit: All segments and boundaries are [+unit]; a blank (in the output from a deletion rule) is [-unit], see Basbøll and Kristensen 1974, p. 220.

Segment: Boundaries are [-segment], all other units are [+segment].

Grammatical boundary: The grammatical boundary (# or /) is [+grammatical boundary], the syllable boundary ($ or ˇ) is [-grammatical boundary]. All segments are unspecified for this

feature, viz. [0 grammatical boundary].[1]

Syllabic: Full and weak vowels are [+syllabic], while obstruents, consonantal sonorants and [j w ɻ ð ɣ] are [-syllabic]. Our general treatment of "syllabicity" is explained in section 2.2 below.

Sonorant: [sonorant] is defined as an acoustic/auditory concept, in agreement with Ladefoged 1971. [p t k f s ʃ h b d g v ʁ ɣ] are [-sonorant], i.e. obstruents, all other segments being [+sonorant].

Constriction: [constriction] is a ternary feature indicating the maximal constriction in the primary speech channel. Nasals and oral stops are [3 constr] (the same applies to trills, taps, etc.). Fricatives are [2 constr], and vocoids, [l] and [h] are [1 constr]. Our use of this ternary feature corresponds to Ladefoged's distinction between stops (i.e. [3 constriction]),

---

1) At the moment, we consider the possibility of changing the grammar so that it can operate with the notion of rank of boundaries (and, hence, rank of rules) as suggested by McCawley, see Basbøll's paper on grammatical boundaries in this volume, particularly pp. 111 f and 119 ff. According to this proposal the binary feature [grammatical boundary] should be replaced by a multivalued feature [boundary], possibly so that $ (or ˇ) is [1 boundary], the intra-word (strong) grammatical boundary (identical with the (weak) inter-word boundary) # (or /) is [2 boundary], the (strong) inter-word boundary # # (or //) is [3 boundary], and the "sentence boundary" (loosely speaking), viz. # # # (or ///) is [4 boundary]. Our notion of $-sensitive versus $-insensitive rules could thus be generalized in such a way that each rule gets its rank specified as [1 bound], [2 bound], etc., by means of an index (1,2,3, or 4) in RULEINDEX, which replaces the present binary distinction of $-sensitive and $-insensitive rules. In the case of a rule of rank 3, boundaries of ranks 1 and 2 (but not boundaries of ranks 3 and 4) should thus be ignored when the compatibility of an input string with the structural description of this rule is examined. From a phonological point of view, this proposed change in our treatment of boundaries seems very attractive. (The fact that zero is compatible with all numbers in our treatment of rule application makes it necessary to keep the distinctive feature [segment], also when the change proposed in this note is carried out.)

fricatives (i.e. [2 constriction]), and approximants (i.e. [1 constriction]); Ladefoged, however, uses the binary features [stop] and [fricative] which permit him to characterize the affricates as [+stop, +fric] (thus in this case ignoring the time dimension). One of the major justifications of the ternary feature [constriction] is that [h] can then be given a reasonable definition, viz. as a voiceless (i.e. [-gl constr]) sound which is [1 constr].[1] It should be noticed that this use of [constriction] permits the inclusion of [1 constr] in the hierarchy of features accounting for the maximal syllabic structure (see Basbøll 1974), at the place between [-consonantal] and [+sonorant], presupposing that [h] is disregarded at the establishment of the hierarchy.[2]

---

1) The ternary feature [constriction] replaces the traditional feature [continuant], [-continuant] being equivalent with [3 constriction]. The distinction between [2 constr] and [1 constr] recalls the distinction between obstruents and sonorants ([2 constr] sounds always being [-sonorant], whereas [1 constr] sounds are normally [+sonorant]), with the important reservation that [h] is [1 constriction] and voiceless, and thus an obstruent, at the same time.

2) [h] is the only Danish segment which is completely non-combinable with any non-syllabic segment, and it thus does not enter into any ordering relations among consonants. This particular status of [h] is codified in the phonological works of Uldall (1936) and Martinet (1937) who consider [h] a prosody (Martinet treats [h] as a breathy attack of (stressed) vowels, which enters the correlation of aspiration: /p, t, k, h/ : /b, d, g, zero/).

The main weakness of our use of the ternary feature [constriction] is that the class of "voiced continuants" cannot be defined as a natural class in the technical sense. This class seems relevant in phonotactics, since the segments which may occur in the vowel-adjacent position in a word-initial three-consonant cluster are [j l ʁ v] (notice that [j] is a non-consonantal sonorant, [l] a consonantal sonorant, and [ʁ v] voiced obstruents). One possible solution is to define this class in terms of two very natural classes, viz. the class of non-syllabic voiced segments minus the class of [3 constriction] (i.e. non-continuant) segments. This is not quite satisfactory, of course.

Lateral: [l] is the only sound which is [+lateral].

Consonantal: [consonantal] is here used as a cover feature (in the sense of Ladefoged), i.e. as a feature which is defined exclusively by means of independently established features. [-consonantal] is defined by the equivalence: [-consonantal] ≡ [+sonorant, 1 constriction, -lateral], and, consequently, the class of all [+consonantal] sounds is the union of the (non-overlapping) classes of obstruents (i.e. [-sonorant]; as mentioned above, all [2 constr] sounds are obstruents), laterals, and non-continuant sonorants (viz. [+sonorant, 3 constr]), i.e. (mainly) nasals. For further details, see Basbøll's paper on diphthongs in this volume, p. 49 ff (as mentioned above, [+continuant] and [1 constriction] are equivalent in the formula).

Glottal constriction: [p t k f s ʃ h ɣ] are [-gl constr], all other segments are [+gl constr]. Notice that [b d g], which phonetically are voiceless (in the traditional sense, i.e. the vocal cords do not vibrate) and distinguished from [p t k] by means of aspiration, are [+gl constr], which agrees well with the results of Frøkjær-Jensen, Ludvigsen and Rischel 1971, as well as with several phonological patterns (in most cases, [gl constr] is identical to the traditional feature [voiced]).

Labial activity: Rounded vowels, including the weakly rounded [ʌ], are [+lab ac]. The same applies to consonants with labial (primary or secondary) articulation, i.e. [w, p, f, b, v, m]. All other phonetic segments are [-lab ac], except [j ɹ h] which are unspecified for this feature.

Apical: [t s d l n ð] are the only sounds which are [+apical].

Back: The normal vowel space is here analyzed by means of two dimensions: [back] and [distance] (measured from the maximally constricted pharyngeal vowel, see below), in addition to [labial activity], as mentioned above (see fig. 1). A vowel is

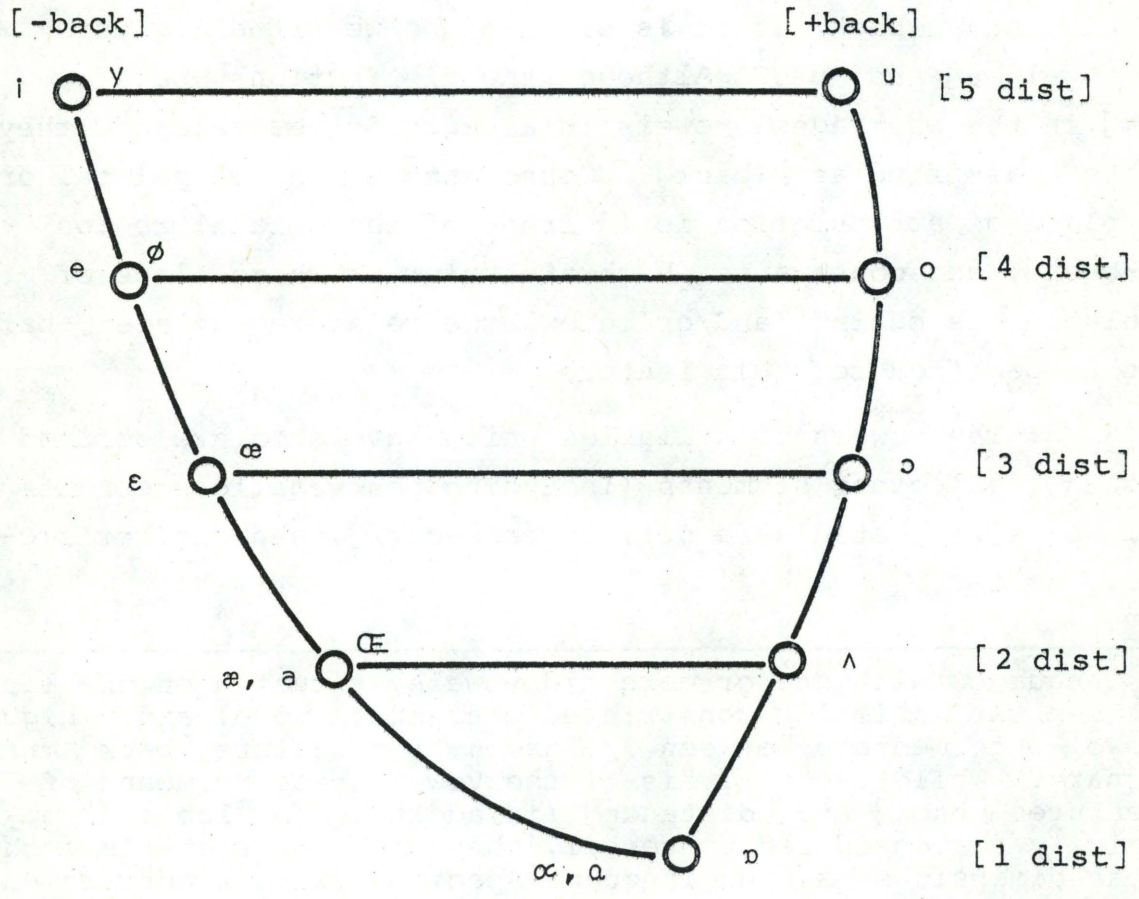

Figure 1

Schematic drawing showing the relationships between the vowels in terms of the features [back] and [distance] (see section 2.1).

[-back] if it is situated on the "left and bottom side" of Jones' Cardinal vowel diagram, i.e. between the highest and most front palatal vowel [i] and the lowest and most back pharyngeal vowel [ɑ], and [+back] if it is situated between the highest velar vowel [u] and [ɒ].[1] Although the distinction [-back]: [+back] in the pharyngeal vowels in a sense is neutralized, they are here classified as [+back]. Consonants which are palatal or whose place of articulation is in front of the palatal region are [-back], and consonants which are velar or whose place of articulation is behind (and/or below) the velar region are [+back]. [h] is unspecified for this feature.[2]

Stød: The vowels of syllables which have stød are defined as [+stød], all other segments (including postvocalic sonorants in syllables with stød) are defined as [-stød], see further section 2.3 below.

---

1) In languages with one or more mid vowels, situated on the line between the maximally constricted pharyngeal vowel and a high mid vowel intermediate between i/y and u, the feature [back] must be ternary. While the analysis of the vowel space by means of the features [back] and [distance] (in addition to [lab ac]) may tentatively be considered universal, the number of steps in each of these dimensions is thus language-specific (within certain universally determined limits, of course).

2) The features [lab ac], [apic] and [back] together seem to cover the traditional dimension "place of articulation". We have, in fact, considered the possibility of operating with a multivalued feature [articulation place] instead, with the coefficients 1 (= labial), 2 (= dental), 3 (= palatal), 4 (= velar), and 5 (= pharyngeal). However, we should still need an independent feature for rounding (cf. the labio-velar glide [w], derived from labi(odent)al [v]). Furthermore, pharyngeal vowels like [ɑ ɒ] sometimes go with the velar vowels (cf. the continuous transition from "maximally velar" to "maximally pharyngeal" vowels), thus constituting a natural class which is unstatable unless one operates with back vowels. The issue is far from settled, however.

Distance: [dist] is a multivalued feature denoting distance from the most constricted pharyngeal vowel. As already mentioned, this feature together with the feature [back] defines two vowel dimensions (see fig. 1); each dimension is partitioned into five steps (i.e. [dist] is a pentavalent feature): [ɑ ɒ ɑ̈] are [1 dist], [æ Œ a ʌ] are [2 dist], [ɛ œ ɔ] are [3 dist], [e ø o] are [4 dist], and [i y u] are [5 dist]. [ɻ ʁ ɣ] are [1 dist], and all other consonants are [5 dist], except [h] which is unspecified for this feature. Our use of the feature [dist] recalls the traditional use of [height] (these features being "inverse proportional", of course). However, [height] is normally considered a "vertical" dimension, which is "perpendicular" on the dimension front-back, with the detrimental consequence that the distinction [a]:[ɑ] is, traditionally, seen purely as one of front:back and not of height. That it is justified to consider e.g. the vowels [i e ɛ æ ɑ] to "lie on the same line", not only physiologically and perceptually, but also phonologically, is shown by the principles of r-colouring which consists of, roughly speaking, decrease by one step in the dimension of [distance].

Aobligatoryshort: The use of the feature [aoblsh] is a "trick", as the name suggests. It is used as the only distinction between the vowels [a ɑ̈], which are [+aoblsh], and [æ ɑ], which are identical to [a ɑ̈], respectively, with the exception that [æ ɑ] are [-aoblsh]. All other segments are unspecified for this feature. The name is due to the fact that [a ɑ̈] only occur as short vowels (when the results of combined ə-assimilation and merger of /VV/ and /V:/ are disregarded, as in da en... [daen/daən/daan/da·n]).[1]

---

1) It may be easier to quantify the output of our grammar (e.g. by turning it into a suitable input to Peter Holtse's speech synthesis) if [aoblsh] is substituted e.g. by a feature "relatively distant", distinguishing [æ ɑ̈] (as [+rel dist]) from [a ɑ] (as [-rel dist]). This change would not affect the rest of our grammar. We have chosen to operate with the feature [aoblsh] since it uses the only obvious phonological difference between the "a-vowels". Phonetically, however, [aoblsh] is even more arbitrary than [rel dist] (which amounts to saying that the least distant unrounded vowels need to be further subdivided with respect to something like distance).

Tense: The feature [tense] distinguishes between full vowels, which are [+tense], and [ɪ ɨ ə], which are [-tense] (the tense counterparts of [ɪ ɨ ə] are [i e ɛ]). All consonants are considered unspecified for this feature.

Grave: [grave] is defined as an acoustic/auditory feature. For the inventory of segments used here, [grave] may be predicted from independent features: all segments which are [+back], as well as all labial consonants, are [+grave]; consequently, all front vowels, as well as all consonants which are neither back nor labial, are [-grave]. ([h] is unspecified for this feature.)

Long: Vowels may be [+long] or [-long], whereas all consonants are considered [-long].

2.2 Syllabicity

As mentioned in the previous section, the full vowels and [ə, ɪ, ɨ] are [+syllabic], whereas all phonetic consonants as well as [j w ɹ ð ɣ] are [-syllabic]. How do we then handle the "schwa-assimilation rules" which create so-called "syllabic consonants", e.g. in handel 'trade' [han?əl] (distinct pronunciation), which is most often pronounced without the vowel [ə], but nevertheless remains a bisyllabic word?

Our grammar contains an optional schwa-deletion rule which changes e.g. [ˇhan?ˇəlˇ] into [ˇhan?ˇlˇ]. Since [l] is the only segment of the second syllable (which is delineated by the two syllable boundaries ˇ), it must be "syllabic", i.e. constitute the peak of this (weak) syllable. This treatment of syllabicity is prosodic.

When [ə] is deleted in a word like kommer 'comes' [ˇkʌm?ˇəɹˇ], the result is [ˇkʌm?ˇɹˇ], i.e. [ɹ] is the peak of the second syllable. This analysis agrees well with the large variability of the syllabic [ˇɹˇ], which may cover the whole range of (unstressed) [ʌ/ɒ].

In a case like kasse 'box' [ˇkasˇəˇ], deletion of [ə] yields the result [ˇkasˇˇ] which suggests a bisyllabic word (where the two-peak-syllabicity may be manifested by length of [s] and by a special intonation). In many cases, at least, where an obstruent should carry the second peak, the bisyllabic word may be reduced to a monosyllable. We may thus operate with an optional rule which deletes a syllable boundary in cases where there is no sonorant adjacent to the deleted schwa.[1]

The syllabic structure thus established (in the notation) by means of the syllable boundaries, together with information on stress (which is not yet available in our grammar), can then be used for the quantification of F_o, intensity and duration which is necessary in the speech synthesis project.

2.3 The stød

We consider the stød to be a prosodic entity, characterizing a syllable with a full vowel as its peak, and indicated as a distinctive feature ([+stød]) of the syllabic peak. Thus only full vowels (which are always [+syllabic, -consonantal, +tense]) can have stød.

It is a consequence of this prosodic treatment of stød that an early stød-rule like "the root-syllable of a prefixed verb gets stød" will assign stød to the short full vowel of words like bekomme 'get', forkaste 'reject', although phonetically there is stød on [m] in the former word and no real stød at all in the latter (see below). It is another consequence of our stød-treatment that the optional vowel shortening rule before

---

1) This formulation, which is meant as a very first approximation, is chosen because it permits syllable reduction regardless of the precise location of the boundary between the syllable from which schwa is deleted, and the preceding syllable. The rule might, for example, delete the former of the two syllable boundaries in the following sequences: ˇ[-son]ˇ, ˇ[-son] [-son]ˇ, and [-son]ˇˇ (as general, the rule should not apply across the grammatical boundary /).

glides need not pay any attention to the location of the stød in a word like ud 'out' [u·?ð, uð?], since the stød will still be a distinctive feature of the vowel also after it has been shortened.

In UNITMATRIX the 16 qualitatively different full vowels occur four times each, viz. as [...., -stød, -long], [...., +stød, -long], [...., -stød, +long], and as [...., +stød, +long] (the order of the 16 vowels is identical in the four groups, each of which may be said to represent a "syllable type"). In the input to the grammar the four syllable types are represented as XVY, XV?Y, XV·Y, and XV·?Y, respectively (unless stød and/or quantity are assigned by rule), where V is a vowel and X and Y arbitrary sequences (including null) of units belonging to the same syllable (including its boundaries).

In the printout, the four syllable types are represented as XVY, XVZ(?)Y (see below), XV·Y, and XV·?Y, respectively. The second type is treated in different ways, according to the unit Z which occurs immediately after the short stød-vowel: if Z is a sonorant or a voiced continuant obstruent, the printout is XVZ?Y; if Z is a voiceless obstruent, or a voiced oral stop, or a boundary, the printout is XVZY, i.e. the stød sign ? is omitted. Thus, in cases like lyst, neuter 'clear (adj.)', or nyhed 'news' [lysd, nyhe·?ð], stød will be dropped in the printout as a consequence of the vowel shortening (of /y·/).

The procedure suggested here, viz. that stød is a characteristic of a syllable with a full vowel, and that it is considered a distinctive feature of the syllabic peak throughout the derivation (but not in the printout), permits the later quantification of F_0, intensity and duration to distinguish between e.g. flæsket as a noun in definite form and as an adjective, a distinction which is found in certain varieties of Standard Danish.

3. Some computational aspects of the project

3.1 Conversion into IPA-notation

As shown in fig. 1 in Basbøll/Kristensen 1974 (p. 218), the overall system was planned to include a subroutine which was to translate the input form in IPA-notation into a string of integers, and another subroutine which was to translate output from strings of integers into forms in IPA-notation. The purpose was to obtain an easy check of the correctness of the input and readability of the output, while internal integer representation of the phonological forms is desirable from a computational point of view. Now, the input form is nothing but a single string, the inventory of units occurring in input forms is only a fraction of all phonetic units, and, furthermore, the corresponding printout includes the input form. The translation can thus be done manually, as far as the input side is concerned (most symbols can be translated directly without recourse to a conversion table, because of the attempted similarity between the IPA-notation of a given input and its representation of characters as an input into the computer). The correspondence between the two character systems: IPA and the keyboard of the data terminal, is secured by a table-specified character conversion on the output side. This character conversion is accomplished in the punching of the output file. The resulting papertape is transferred to printout in IPA-notation by means of a papertape typewriter equipped with an IPA ball-head. (It should be remembered that the MAIN PROGRAM itself contains subroutines which translate the keyboard representation of a string into integer representation and back again.)

3.2 Description of a RUN

The rule testing program is run via a UNISCOPE 100 demand terminal. Using the ED processor, one may update RULEINDEX and RULEMATRIX and insert one or more input strings into a temporary

file (the manual transformation from input in IPA-notation to the keyboard of the scope is quite simple, cf. above). The MAIN PROGRAM is stored in a permanent file in its symbolic form and in its absolute form. The program works on four types of data: 1) UNITMATRIX which is stored in the permanent file; 2) RULE-INDEX and RULEMATRIX which, too, are stored in the permanent file; 3) rulelimits (see secion 3.4 below) are normally put in directly via the keyboard; 4) the input string(s) which may be stored in a temporary or permanent file[1] or put in directly.

Each RUN has two output files: 1) a file with printout consisting of RULEINDEX and RULEMATRIX which may be led to a line printer; 2) a punch tape file containing the input and output strings to and from the rules of the grammar, together with the designations H and L[2]. A printout of UNITMATRIX can be obtained

---

1) At the moment, the data used for the phonotactic surveys of Jespersen 1926 and Vestergaard 1968 is stored in some file elements, which is expedient from the point of view of testing of rules (e.g. syllabification rules). We plan to store much more material of this kind. It is, for example, our intention to cover systematically all different /rV/-sequences as well as all different /VC/-sequences in which /V/ can be part of the relevant context for structural changes in /C/, and conversely.

2) In Basbøll/Kristensen 1974 (p. 225) we proposed to use the designations A (meaning "the obligatory rule was applied non-vacuously"), V (meaning "the rule was applied vacuously"), and O (meaning "application of the rule was tried, but its structural description was not satisfied"), in addition to L (meaning "the optional rule was applied non-vacuously") and H (meaning "non-application of an optional rule, the application of which would give rise to an L-form"). Since we no longer consider the information offered by the designations O and V phonologically important, we have omitted O and V. Furthermore, the information represented by an A is completely redundant in the printout since the rules which have been non-vacuously applied are listed, and if such a rule has no L-designation the rule must be obligatory, i.e. A. The only designations kept in the printout are, therefore, H and L (or, with the characters of the IPA ball-head, h and l), which saves considerable punching time and paper space.

after insertion of a couple of WRITE-statements into the MAIN PROGRAM. This seems inelegant, but updating of UNITMATRIX is rather rare.

The bipartition of the output from the data processing (cf. 1) and 2) above) is desirable because RULEINDEX and RULEMATRIX mainly consist of integers which the IPA ball-head does not contain, while the writing chain of the line printer does not contain the IPA-symbols. (Later on when updating of RULEINDEX and RULEMATRIX will not be necessary any more, the WRITE-statements causing these to be printed out can be omitted, and the system will be simpler.)

3.3 Optional rules

In Basbøll/Kristensen 1974, p. 223-224, it was described how the MAIN PROGRAM was to handle the facultativity which is implicated by the optional rules of the grammar: If the grammar contains n optional rules, each input form was to follow 2^n different paths of derivation. This method has now proved to be clearly uneconomic, for most often only a few of the optional rules are relevant to a given input string (as we realized, in fact, on p. 224!). In a test of a subcomponent of the grammar containing about 10 optional rules, the RUNtimes were found to be excessively long. A new structure of the MAIN PROGRAM was painfully necessary. Now a minimum of different paths of derivation are followed. First of all, the input string is taken through the grammar. Hereby the program tries to apply all the rules of the grammar, optional or not, to the output form from the preceding rule (or, in case, the input form to the grammar). Every time the input string satisfies the structural description of an optional rule and this input string is changed by the application of the rule (i.e. the application is non-vacuous), a node is established. When all the rules have been run through, and the output forms and the designations H and L have been transferred to the output file, the last established node is taken from the node

list together with the number of the (optional) rule where the node was established. From this point of the grammar all L paths with respect to this rule have already been followed. The H paths with respect to this rule have the same derivational history as the L paths before the optional rule in question, and the different variables still contain the relevant output forms and designations. Now the program deletes the node in question and goes on to the next rule (without trying to apply the optional rule whose node has just been deleted), application of the remaining rules is tried, and maybe new nodes are established. This procedure is continued until the node list is empty. The method is economic because every time a new node is handled, the computer processes data already in existence, placing a new layer of data over data already written out. The gain from this revision of the MAIN PROGRAM is really considerable.

3.4 Rule limits

As hinted at in section 3.2, it is now possible for us to test arbitrary parts of the grammar. This is desirable from an economic point of view, because the RUNtime will be shortened (in particular when only obligatory rules are tested), and also because a lot of information which is irrelevant for the problem at hand can be avoided. The part of the grammar to be tested is selected by deciding the limits in terms of rule numbers for the block or blocks of rules which one wants to have included in the desired grammar. The limits ordered in sets of pairs of (rule) numbers are input to the data processing, as mentioned in section 3.2. Moreover, such sets of rule limits corresponding to different parts of the grammar (e.g. those rule limits defining the set of all obligatory rules, giving rise to the most distinct and conservative output form of each input form) may be stored in a permanent file and added as input in a single command.

```
input  ///indspʁøʔjt///
i
   output fʁɑ ʁegel nʁ
   i          //ˇ/indspʁøj?t///
   v          //ˇ/indspʁøj?tˇ///
   xix        //ˇ/indˇspʁøj?tˇ///
   xxvii      //ˇ/endˇspʁøj?tˇ///
   xxix       //ˇ/enðˇspʁøj?tˇ///
   xxxi       //ˇ/enˇspʁøj?tˇ///
   xxxv       //ˇ/enˇspʁʌj?tˇ///
   lxi        //ˇ/enˇsbʁʌj?tˇ///
input  ///stegning///
i
   output fʁɑ ʁegel nʁ
   iii        //ˇ/stegning///
   v          //ˇ/stegningˇ///
   xiv        //ˇ/stegˇningˇ///
   xxii       //ˇ/stegˇniŋgˇ///
   xxvii      //ˇ/stegˇneŋgˇ///
   xxix       //ˇ/steɣˇneŋɣˇ///
   xxxii      //ˇ/steɣˇneŋˇ///
   xxxiv      //ˇ/stajˇneŋˇ///
   lxi        //ˇ/sdajˇneŋˇ///
input  ///ty?ngdə///
i
   output fʁɑ ʁegel nʁ
   ii         //ˇ/tyn?gdə///
   v          //ˇ/tyn?gdəˇ///
   xv         //ˇ/tyn?gˇdəˇ///
   xxii       //ˇ/tyŋ?gˇdəˇ///
   xxvii      //ˇ/tøŋ?gˇdəˇ///
   xxix       //ˇ/tøŋ?ɣˇdəˇ///
   xxxii      //ˇ/tøŋ?ˇdəˇ///
input  ///ʁɛdsɔ?m///
i
   output fʁɑ ʁegel nʁ
   ii         //ˇ/ʁɛdsɔm?///
   v          //ˇ/ʁɛdsɔm?ˇ///
   xiv        //ˇ/ʁɛdˇsɔm?ˇ///
   xxix       //ˇ/ʁɛðˇsɔm?ˇ///
   xxxix      //ˇ/ʁæðˇsɔm?ˇ///
   xliv       //ˇ/ʁæðˇsʌm?ˇ///
input  ///fængəhol///
i
   output fʁɑ ʁegel nʁ
   ii         //ˇ/fængəhol///
   v          //ˇ/fængəholˇ///
   viii       //ˇ/fængəˇholˇ///
   ix         //ˇ/fængˇəˇholˇ///
   xxii       //ˇ/fæŋgˇəˇholˇ///
   xxix       //ˇ/fæŋɣˇəˇholˇ///
   xxxii      //ˇ/fæŋˇəˇholˇ///
   xxxiii     //ˇ/faŋˇəˇholˇ///
   xlv        //ˇ/faŋˇəˇhɔlˇ///
   xlvi       //ˇ/fʁŋˇəˇhɔlˇ///
```

Figure 2

Derivations of the words ind-sprøjt, stegning, tyngde, ræd-som, and fangehul. Only obli-gatory rules are applied, and each word has, therefore, only one version (see section 4).

```
input  ///kʁæ·ʔv///
i
   output fʁɑ ʁegel nʁ
   iii        //ˇ/kʁæ·ʔv///
   v          //ˇ/kʁæ·ʔvˇ///
   xxxviii    //ˇ/kʁɑ·ʔvˇ///
   lv         //ˇ/kʁɑ·ʔwˇ///
   lxxi       //ˇ/kʁɑwʔˇ///
   ʁegel      bem
   lv          l
   lxxi        l
ii
   output fʁɑ ʁegel nʁ
   iii        //ˇ/kʁæ·ʔv///
   v          //ˇ/kʁæ·ʔvˇ///
   xxxviii    //ˇ/kʁɑ·ʔvˇ///
   lv         //ˇ/kʁɑ·ʔwˇ///
   ʁegel      bem
   lv          l
   lxxi        h
iii
   output fʁɑ ʁegel nʁ
   iii        //ˇ/kʁæ·ʔv///
   v          //ˇ/kʁæ·ʔvˇ///
   xxxviii    //ˇ/kʁɑ·ʔvˇ///
   ʁegel      bem
   lv          h
```

```
input  ///flæ·ʔd///
i
   output fʁɑ ʁegel nʁ
   iii        //ˇ/flæ·ʔd///
   v          //ˇ/flæ·ʔdˇ///
   xxix       //ˇ/flæ·ʔðˇ///
   lxxi       //ˇ/flæðʔˇ///
   lxxii      //ˇ/flɑðʔˇ///
   ʁegel      bem
   lxxi        l
   lxxii       l
ii
   output fʁɑ ʁegel nʁ
   iii        //ˇ/flæ·ʔd///
   v          //ˇ/flæ·ʔdˇ///
   xxix       //ˇ/flæ·ʔðˇ///
   lxxi       //ˇ/flæðʔˇ///
   ʁegel      bem
   lxxi        l
   lxxii       h
iii
   output fʁɑ ʁegel nʁ
   iii        //ˇ/flæ·ʔd///
   v          //ˇ/flæ·ʔdˇ///
   xxix       //ˇ/flæ·ʔðˇ///
   ʁegel      bem
   lxxi        h
```

Figure 3

Derivations of the words krav and flad. Both optional and obligatory rules are applied. Each word has three versions (see section 4).

input ///kʁi•ʔg///

i

output	fʁɑ	ʁegel	nʁ
iii	//ˇ/kʁi•ʔg///		
v	//ˇ/kʁi•ʔgˇ///		
xxix	//ˇ/kʁi•ʔɣˇ///		
l	//ˇ/kʁi•ʔjˇ///		
lxviii	//ˇ/kʁi•ʔˇ///		
ʁegel	bem		
l	l		
lxviii	l		

ii

output	fʁɑ	ʁegel	nʁ
iii	//ˇ/kʁi•ʔg///		
v	//ˇ/kʁi•ʔgˇ///		
xxix	//ˇ/kʁi•ʔɣˇ///		
l	//ˇ/kʁi•ʔjˇ///		
lxxi	//ˇ/kʁijʔˇ///		
ʁegel	bem		
l	l		
lxviii	h		
lxxi	l		

iii

output	fʁɑ	ʁegel	nʁ
iii	//ˇ/kʁi•ʔg///		
v	//ˇ/kʁi•ʔgˇ///		
xxix	//ˇ/kʁi•ʔɣˇ///		
l	//ˇ/kʁi•ʔjˇ///		
ʁegel	bem		
l	l		
lxviii	h		
lxxi	h		

iv

output	fʁɑ	ʁegel	nʁ
iii	//ˇ/kʁi•ʔg///		
v	//ˇ/kʁi•ʔgˇ///		
xxix	//ˇ/kʁi•ʔɣˇ///		
lxviii	//ˇ/kʁi•ʔˇ///		
ʁegel	bem		
l	h		
lxviii	l		

v

output	fʁɑ	ʁegel	nʁ
iii	//ˇ/kʁi•ʔg///		
v	//ˇ/kʁi•ʔgˇ///		
xxix	//ˇ/kʁi•ʔɣˇ///		
lxxi	//ˇ/kʁiɣʔˇ///		
ʁegel	bem		
l	h		
lxviii	h		
lxxi	l		

vi

output	fʁɑ	ʁegel	nʁ
iii	//ˇ/kʁi•ʔg///		
v	//ˇ/kʁi•ʔgˇ///		
xxix	//ˇ/kʁi•ʔɣˇ///		
ʁegel	bem		
l	h		
lxviii	h		
lxxi	h		

Figure 4

Derivations of the word krig. Both optional and obligatory rules are applied. The word has six versions (see section 4).

```
input  ///spʁing/u·ʔd//idio·ʔt///
i
   output fʁɑ ʁegel nʁ
   i          ///spʁingˇ/u·ʔd/ˇ/idio·ʔt///
   iv         //ˇ/spʁingˇ/u·ʔd/ˇ/idio·ʔt///
   v          //ˇ/spʁingˇ/u·ʔd/ˇ/idio·ʔtˇ///
   vi         //ˇ/spʁingˇ/u·ʔd/ˇ/idiˇo·ʔtˇ///
   viii       //ˇ/spʁingˇ/u·ʔd/ˇ/iˇdiˇo·ʔtˇ///
   xxii       //ˇ/spʁiŋgˇ/u·ʔd/ˇ/iˇdiˇo·ʔtˇ///
   xxvii      //ˇ/spʁeŋgˇ/u·ʔd/ˇ/iˇdiˇo·ʔtˇ///
   xxix       //ˇ/spʁeŋɣˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   xxxii      //ˇ/spʁeŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   xliii      //ˇ/spʁɛŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   lxi        //ˇ/sbʁɛŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   lxvi       //ˇ/sbʁɛŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔdˇ///
   lxxi       //ˇ/sbʁɛŋˇ/uðʔ/ˇ/iˇdiˇo·ʔdˇ///
   ʁegel      bem
   xliii       l
   lxvi        l
   lxxi        l
ii
   output fʁɑ ʁegel nʁ
   i          ///spʁingˇ/u·ʔd/ˇ/idio·ʔt///
   iv         //ˇ/spʁingˇ/u·ʔd/ˇ/idio·ʔt///
   v          //ˇ/spʁingˇ/u·ʔd/ˇ/idio·ʔtˇ///
   vi         //ˇ/spʁingˇ/u·ʔd/ˇ/idiˇo·ʔtˇ///
   viii       //ˇ/spʁingˇ/u·ʔd/ˇ/iˇdiˇo·ʔtˇ///
   xxii       //ˇ/spʁiŋgˇ/u·ʔd/ˇ/iˇdiˇo·ʔtˇ///
   xxvii      //ˇ/spʁeŋgˇ/u·ʔd/ˇ/iˇdiˇo·ʔtˇ///
   xxix       //ˇ/spʁeŋɣˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   xxxii      //ˇ/spʁeŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   xliii      //ˇ/spʁɛŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   lxi        //ˇ/sbʁɛŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   lxvi       //ˇ/sbʁɛŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔdˇ///
   ʁegel      bem
   xliii       l
   lxvi        l
   lxxi        h
iii
   output fʁɑ ʁegel nʁ
   i          ///spʁingˇ/u·ʔd/ˇ/idio·ʔt///
   iv         //ˇ/spʁingˇ/u·ʔd/ˇ/idio·ʔt///
   v          //ˇ/spʁingˇ/u·ʔd/ˇ/idio·ʔtˇ///
   vi         //ˇ/spʁingˇ/u·ʔd/ˇ/idiˇo·ʔtˇ///
   viii       //ˇ/spʁingˇ/u·ʔd/ˇ/iˇdiˇo·ʔtˇ///
   xxii       //ˇ/spʁiŋgˇ/u·ʔd/ˇ/iˇdiˇo·ʔtˇ///
   xxvii      //ˇ/spʁeŋgˇ/u·ʔd/ˇ/iˇdiˇo·ʔtˇ///
   xxix       //ˇ/spʁeŋɣˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   xxxii      //ˇ/spʁeŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   xliii      //ˇ/spʁɛŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   lxi        //ˇ/sbʁɛŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   lxxi       //ˇ/sbʁɛŋˇ/uðʔ/ˇ/iˇdiˇo·ʔtˇ///
   ʁegel      bem
   xliii       l
   lxvi        h
   lxxi        l
iv
   output fʁɑ ʁegel nʁ
   i          ///spʁingˇ/u·ʔd/ˇ/idio·ʔt///
   iv         //ˇ/spʁingˇ/u·ʔd/ˇ/idio·ʔt///
   v          //ˇ/spʁingˇ/u·ʔd/ˇ/idio·ʔtˇ///
   vi         //ˇ/spʁingˇ/u·ʔd/ˇ/idiˇo·ʔtˇ///
   viii       //ˇ/spʁingˇ/u·ʔd/ˇ/iˇdiˇo·ʔtˇ///
   xxii       //ˇ/spʁiŋgˇ/u·ʔd/ˇ/iˇdiˇo·ʔtˇ///
   xxvii      //ˇ/spʁeŋgˇ/u·ʔd/ˇ/iˇdiˇo·ʔtˇ///
   xxix       //ˇ/spʁeŋɣˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   xxxii      //ˇ/spʁeŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   xliii      //ˇ/spʁɛŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   lxi        //ˇ/sbʁɛŋˇ/u·ʔð/ˇ/iˇdiˇo·ʔtˇ///
   ʁegel      bem
   xliii       l
   lxvi        h
   lxxi        h
```

Figure 5 (part 1)

v

output	fʁɑ ʁegel nʁ
i	///spʁingˇ/u·?d/ˇ/idio·?t///
iv	//ˇ/spʁingˇ/u·?d/ˇ/idio·?t///
v	//ˇ/spʁingˇ/u·?d/ˇ/idio·?tˇ///
vi	//ˇ/spʁingˇ/u·?d/ˇ/idiˇo·?tˇ///
viii	//ˇ/spʁingˇ/u·?d/ˇ/iˇdiˇo·?tˇ///
xxii	//ˇ/spʁiŋgˇ/u·?d/ˇ/iˇdiˇo·?tˇ///
xxvii	//ˇ/spʁɐŋgˇ/u·?d/ˇ/iˇdiˇo·?tˇ///
xxix	//ˇ/spʁɐɣˇ/u·?ð/ˇ/iˇdiˇo·?tˇ///
xxxii	//ˇ/spʁɐɳˇ/u·?ð/ˇ/iˇdiˇo·?tˇ///
lxi	//ˇ/sbʁɐɳˇ/u·?ð/ˇ/iˇdiˇo·?tˇ///
lxvi	//ˇ/sbʁɐɳˇ/u·?ð/ˇ/iˇdiˇo·?dˇ///
lxxi	//ˇ/sbʁɐɳˇ/uð?/ˇ/iˇdiˇo·?dˇ///
ʁegel	bem
xliii	h
lxvi	l
lxxi	l

vi

output	fʁɑ ʁegel nʁ
i	///spʁingˇ/u·?d/ˇ/idio·?t///
iv	//ˇ/spʁingˇ/u·?d/ˇ/idio·?t///
v	//ˇ/spʁingˇ/u·?d/ˇ/idio·?tˇ///
vi	//ˇ/spʁingˇ/u·?d/ˇ/idiˇo·?tˇ///
viii	//ˇ/spʁingˇ/u·?d/ˇ/iˇdiˇo·?tˇ///
xxii	//ˇ/spʁiŋgˇ/u·?d/ˇ/iˇdiˇo·?tˇ///
xxvii	//ˇ/spʁɐŋgˇ/u·?d/ˇ/iˇdiˇo·?tˇ///
xxix	//ˇ/spʁɐɣˇ/u·?ð/ˇ/iˇdiˇo·?tˇ///
xxxii	//ˇ/spʁɐɳˇ/u·?ð/ˇ/iˇdiˇo·?tˇ///
lxi	//ˇ/sbʁɐɳˇ/u·?ð/ˇ/iˇdiˇo·?tˇ///
lxvi	//ˇ/sbʁɐɳˇ/u·?ð/ˇ/iˇdiˇo·?dˇ///
ʁegel	bem
xliii	h
lxvi	l
lxxi	h

vii

output	fʁɑ ʁegel nʁ
i	///spʁingˇ/u·?d/ˇ/idio·?t///
iv	//ˇ/spʁingˇ/u·?d/ˇ/idio·?t///
v	//ˇ/spʁingˇ/u·?d/ˇ/idio·?tˇ///
vi	//ˇ/spʁingˇ/u·?d/ˇ/idiˇo·?tˇ///
viii	//ˇ/spʁingˇ/u·?d/ˇ/iˇdiˇo·?tˇ///
xxii	//ˇ/spʁiŋgˇ/u·?d/ˇ/iˇdiˇo·?tˇ///
xxvii	//ˇ/spʁɐŋgˇ/u·?d/ˇ/iˇdiˇo·?tˇ///
xxix	//ˇ/spʁɐɣˇ/u·?ð/ˇ/iˇdiˇo·?tˇ///
xxxii	//ˇ/spʁɐɳˇ/u·?ð/ˇ/iˇdiˇo·?tˇ///
lxi	//ˇ/sbʁɐɳˇ/u·?ð/ˇ/iˇdiˇo·?tˇ///
lxxi	//ˇ/sbʁɐɳˇ/uð?/ˇ/iˇdiˇo·?tˇ///
ʁegel	bem
xliii	h
lxvi	h
lxxi	l

viii

output	fʁɑ ʁegel nʁ
i	///spʁingˇ/u·?d/ˇ/idio·?t///
iv	//ˇ/spʁingˇ/u·?d/ˇ/idio·?t///
v	//ˇ/spʁingˇ/u·?d/ˇ/idio·?tˇ///
vi	//ˇ/spʁingˇ/u·?d/ˇ/idiˇo·?tˇ///
viii	//ˇ/spʁingˇ/u·?d/ˇ/iˇdiˇo·?tˇ///
xxii	//ˇ/spʁiŋgˇ/u·?d/ˇ/iˇdiˇo·?tˇ///
xxvii	//ˇ/spʁɐŋgˇ/u·?d/ˇ/iˇdiˇo·?tˇ///
xxix	//ˇ/spʁɐɣˇ/u·?ð/ˇ/iˇdiˇo·?tˇ///
xxxii	//ˇ/spʁɐɳˇ/u·?ð/ˇ/iˇdiˇo·?tˇ///
lxi	//ˇ/sbʁɐɳˇ/u·?ð/ˇ/iˇdiˇo·?tˇ///
ʁegel	bem
xliii	h
lxvi	h
lxxi	h

Figure 5 (part 2)

Derivations of the utterance Spring ud, idiot! Both optional and obligatory rules are applied. The utterance has eight versions (see section 4).

4. Examples of derivations

Before we conclude this report (in section 5) by stating some areas which our project might be enlargened to cover, we shall refer the reader to fig.s 2-5 which contain examples of derivations within the present version of our grammar.

Fig. 2 contains derivations of the words indsprøjt 'inject', stegning 'roasting',[1] tyngde 'heaviness', rædsom 'horrible', and fangehul 'dungeon'; only the obligatory rules are applied, and the output should thus be distinct, conservative pronunciations (but only in cases where alternative pronunciations exist, of course).

Fig. 3 contains derivations of the words krav 'demand' and flad 'flat', where all rules, including the optional ones, are included in the testing. Each word has three different versions.

Fig. 4 contains derivations of the word krig 'war', all rules being included in the testing. The two outputs [kʁɐˑʔ] are both included in the printout since they have different derivational histories (corresponding to the phonological fact that the final non-syllabic segment may be dropped after high vowels, both in standards with [ɣ] as a separate segment, and in younger standards where [ɣ] has been replaced by [j, w]).

Fig. 5 contains the derivations of the utterance spring ud, idiot! 'jump, (you) idiot!'. There are eight versions of this utterance, corresponding to the optionality of r-colouring of e,

---

1) We have used the orthographic form /stegning/ as input, in agreement with section 5 (ii) below, although the full vowel /i/ ought, from a purely phonological point of view, to be substituted by the lax vowel /ə/ (see section 2 above). In that case, the derivation of the second syllable would have passed through the derivational stages nəŋg , nɪŋg (by vowel raising of schwa before velars), and niŋg (by lowering of high front vowels before nasals).

of vowel shortening before ð, and of de-aspiration of an utterance-final plosive (notice that these phenomena are independent of each other, in contradistinction to the optional rules applied to krav, flad in fig. 3).

5. Further work

There are several directions into which we may continue our project. Three of these will be mentioned below (there are several others, e.g. concerning an automatic determination of the redundancy of UNITMATRIX, which will not be discussed here).

(i) We try to make the phonetic output of our rules so specific, detailed, and phonetically realistic that it can be used as input to Peter Holtse's project of speech synthesis by rule of Standard Danish. One aspect of this coordination is the attempt to use a phonetically satisfying distinctive feature analysis of our units, although it may sometimes seem too redundant and unelegant from an abstract phonological point of view (cf. section 2.1 above). Our cooperation with Peter Holtse is planned to continue.

(ii) Concerning the more abstract parts of our grammar, we try to approach the possibility of using orthographic forms as input to our rule system (as can be seen from the examples of derivations in fig.s 2-5, our input forms are at present mostly very close to orthographic forms, the main deviation being that we need, so far, more information than the writing gives as to stød and the distinction of schwa vs. the full vowel /e/ the standard writing provides us with). Our attempt to use orthographic or near-orthographic forms as input to the greatest

possible extent,[1] is justified for at least two reasons: firstly, the input will be well-defined and not "open-ended", cf. the great difficulties in giving a non-arbitrary characterization of the systematic phonemic representations in generative phonology; and, secondly, we may be able to change writing into speech (by including the work suggested under (i) above), which opens up wide perspectives of practical use.

(iii) The third main line of our project is one of using the grammar we have constructed to investigate the notion of variable (or optional) rules, e.g. as to their possible interrelationship in a hierarchical (or other) structure. Thus the output forms of our grammar may be spoken by a person on tape (or, according to (i) above, may be realized by means of speech synthesis), and then evaluated by a number of informants as to acceptability, stylistic value, etc. Such an investigation might shed light on important issues in synchronic grammar, concerning the real nature of speech variation and variable rules.

---

1) It is very simple to change the input to the grammar into "quasi-orthographic forms" by means of "rewrite-rules", viz. rules which double a single consonant between a short vowel and schwa, omit ?, rewrite ɛ, æ, ə, ɔ as æ, a, e and aa, etc. Such a quasi-orthographic notation immediately reveals the points where there is a non-superficial discrepancy between standard writing and our phonological forms. This might be of use in dealing with orthographical issues, since one may thereby find cases in which orthography could be made more regular (there are, of course, other kinds of information, such as spelling errors, which are more important in that context).

References

Basbøll, Hans 1974: "The phonological syllable with special reference to Danish", ARIPUC 8, p. 39-128

Basbøll, Hans and Kjeld Kristensen 1974: "Preliminary work on computer testing of a generative phonology of Danish", ARIPUC 8, p. 216-226

Frøkjær-Jensen, Børge, Carl Ludvigsen and Jørgen Rischel 1971: "A glottographic study of some Danish consonants", Form & Substance (L.L. Hammerich, R. Jakobson and E. Zwirner, eds.) (Copenhagen), p. 123-140 (also in ARIPUC 7 (1973), p. 269-295)

Jespersen, Otto 1926: Modersmålets fonetik (Copenhagen)

Martinet, André 1937: "La phonologie du mot en danois", BSL 38, p. 169-266 (also published as an independent book)

McCawley, James D. 1968: The phonological component of a grammar of Japanese (The Hague)

Ladefoged, Peter 1971: Preliminaries to linguistic phonetics (Chicago)

Uldall, Hans Jørgen 1936: "The phonematics of Danish", Proc. Phon. 2, p. 54-57

Vestergaard, Torben 1968: "Initial and final consonant combinations of Danish monosyllables", SL 21, p. 37-66