SUMMARIES OF PH.D. THESES DEFENDED IN 1998

VOICE CHARACTERISTICS FOLLOWING RADIOTHERAPY: THE DEVELOPMENT OF A PROTOCOL

author: Irma M. Verdonck-de Leeuw
promotor: Louis C.W. Pols
co-promotor: Florien J. Koopmans-van Beinum
date of defence: February 3, 1998

Summary

Prognosis concerning survival is good for patients who are treated with radiotherapy for early glottic cancer, with cure rates of 70-90%. Despite these good results, there is still uncertainty about the optimal radiation dose. The optimal dose should be based on tumour control and possible complications. Voice worsening can be a complication of radiotherapy. This thesis aims at some of the theoretical, practical, and methodological problems of voice analyses in order to assess possible outcomes of radiotherapy on voice characteristics in terms of voice quality, vocal function, and vocal performance.

A literature survey (Chapter 1) reveals that few studies are carried out on voice characteristics of patients following radiotherapy for early glottic cancer. In addition, results of the 19 studies reviewed are hard to compare because of methodological differences. Most striking is the variety of speakers: men and women ranging in age, with small to large tumours, treated with different radiation schedules, before, during, and right after radiation up to ten years after radiotherapy. Therefore, it is striking too that only in six studies control speakers were involved. In the other studies, patient groups were compared with themselves at various moments before and after treatment or with mean data from the literature. Furthermore, several voice analyses are applied: perceptual voice ratings, acoustical voice measurements, or clinical methods such as phonetography and stroboscopy. Although it is hard to compare results of these studies, it can be concluded that an acute effect of radiotherapy on voice characteristics has been shown, but that late effects are still obscure.

Before examining this, a description is given in Chapter 2 of the "normal" anatomy and physiology of the larynx, of early glottic cancer, and of the treatment this thesis focuses on: radiotherapy. Also, the trial study is described, that is carried out at the Netherlands Cancer Institute/Antoni van Leeuwenhoekhuis and that deals

with the effect of two different radiation schedules for early glottic carcinoma; this thesis is part of that trial study.

Chapter 3 comprises a detailed description of the 60 patients and 20 control speakers who have participated in this research project. Because voice characteristics are speaker dependent, a group of ten patients is followed from before radiation, six months after up to two years after radiotherapy (n=30). Further follow-up of these patients fell out of the range of the project, but because possible late effects should become visible or audible as well, five separate groups of patients were composed: before radiation, six months after, two years after, three to seven years after, and seven to ten years after rae project; these speakers were matched with the patients concerning sex (all male), age (between 51 and 81 years old), and smoking and drinking habits. The group arrangement is applied to develop a protocol of voice analyses, in the course of which it is investigated which analyses can differentiate these speaker groups best. Subsequently, voice characteristics following radiotherapy are examined even more precisely, dependent on five aspects: stage of the tumour (unilateral or bilateral), initial surgery (biopsying or stripping the vocal fold), radiation schedule (66 Gy in 33 fractions, 60 Gy in 30 fractions, or 60 Gy in 25 fractions), age of the speaker (younger than 65 years, between 65 and 70 years, between 70 and 75 years, or older than 75 years), and whether or not smoking was continued after treatment. But before these aspects are discussed, first a description is given of the development of the protocol concerning perceptual analyses of voice quality (Chapter 4), different pitch analyses (Chapter 5), and acoustical analyses of voice quality (Chapter 6).

Chapter 4 deals with perceptual analyses of voice quality. Ratings from three trained and 20 naive raters and from the speakers themselves and their partners are gathered. The trained raters are trained in the use of the 'Vocal Profile Analysis Protocol' by John Laver; the naive raters and the speakers themselves and their partners judge voice quality on seven-points scales that are especially developed for naive Dutch raters. The trained and naive raters judge voice quality on read-aloud text and on sustained /a/ vowels. Trained raters are found to be more reliable than naive raters, but reliability is satisfactory for both rater groups; reliability could neither be assessed for the ratings of the speakers themselves nor for their partners, since they rated just one voice at the time. Furthermore it appears that patients before radiotherapy have the most deviant voice quality; voice quality of patients six months, two years, and three to seven years after radiation is less deviant, but still significantly worse than voice quality of the control speakers; patients seven to ten years after radiotherapy are comparable with control speakers. This trend is found most obviously for the trained raters on read-aloud text on the scales breathiness, roughness, and tension. The conclusion is that perceptual analysis of voice quality by trained raters is preferred.

It would seem that voice quality can be analysed by means of perceptual judgements. However, there are still certain shortcomings attached to this method. Even though reliability of the raters has been shown, their ratings remain subjective. Furthermore, perceptual analyses are very time-consuming, which is a considerable drawback, especially in clinical practice. Sufficient reason to draw the attention to acoustical analyses of voice quality, which are objective and quick to perform. In Chapter 5, a closer look is taken at pitch analysis. Perceptual, acoustical, and electroglottographic analyses are compared. Earlier research revealed that perceptual pitch ratings may be influenced by deviant voice quality. Acoustical analyses of fundamental frequency (pitch is the audible feature we attach to differences in fundamental frequency) are probably less disturbed by deviant voice quality. However, acoustic signals do contain strong harmonics due to the resonant

frequencies of the vocal tract (oral/pharynx cavity) which may hamper 'pitch extraction'. Electroglottographic (EGG) signals represent vocal fold activity (and thereby fundamental frequency) more directly and are therefore taken into account to determine which method can best be used to analyse pitch of pathological voices. Results show that perceptual analyses are indeed influenced by deviant voice quality. Raters have problems particularly with rough voices: these are often judged as lower, while they are not that low. Results from the objective acoustic and electroglotto-graphic analyses are comparable, provided that the analyses are well performed. Nevertheless, preference is given to acoustical pitch analysis, because no reliable EGG-signals could be obtained from more than 20% of the speakers.

In Chapter 6, acoustical analyses of voice quality are further examined. By means of the speech processing system PRAAT developed by Boersma (Institute of Phonetic Sciences) the mean fundamental frequency and the harmonics-to-noise ratio are analysed. Besides that, the commercially available package Multidimensional Voice Program (MDVP) provides a series of parameters that are grouped under fundamental frequency, frequency and amplitude perturbation (jitter and shimmer), voice breaks, voice irregularities, noise, and tremor. Finally, a new parameter is used: duration of voice onset of the sustained /a/; this is measured manually. Again, results are compared with perceptual ratings (breathiness, roughness, and tension) by trained and naive raters on read-aloud text and the sustained /a/, to determine which analyses can best be used. It appears that acoustical analyses (especially standard deviation of the fundamental frequency, jitter, noise, and duration of the voice onset) show the same trend as was found for the perceptual ratings, albeit less strong. Direct single correlations between acoustical and perceptual voice parameters are low; results of multiple regression analyses show that a perceptual parameter can be predicted better by a set of acoustical measures. The conclusion is that, in the case of separate speaker groups, voice quality can best be analysed by means of scale judgements by trained raters. For a longitudinal research design, acoustical measures are objective and quick to perform and come close to judgements by naive raters.

Besides analyses of voice quality, measures of vocal function are also of interest in investigating the effect of radiotherapy on voice characteristics. In Chapter 7 the phonetogram, maximum phonation time, phonation quotient, and evaluations of video-laryngo-stroboscopy are used to investigate vocal function. It appears that frequency and amplitude range, measured by means of phonetography, maximum phonation time, and phonation quotient give insufficient insight into vocal function following radiotherapy. These measures are left aside. Stroboscopy, on the other hand, although unpleasant for the speaker and therefore not available for all speakers, does give a lot of information. It appears that patients after radiotherapy have more glottic oedema and more vascular injection on the vocal fold and that the vocal fold edge is often irregular, that the mucosal wave is often diminished, that a nonvibrating portion of the vocal fold is often present, and that vocal fold closure is often incomplete. Furthermore, it appears that in addition to increasing age of the speaker and stripping instead of biopsying the vocal fold (which was also found to have an adverse effect for perceptual analyses of voice quality), also continuing smoking after radiotherapy decrease vocal function.

In Chapter 8 the effect of a voice disorder on daily life is investigated. The speakers are asked to indicate their vocal performance by means of self-ratings on several scales, such as the ability to shout, have a normal (telephone) conversation, the amount of getting tired from speaking, and the avoidance of a large party. Their answers were compared with the earlier derived measures for voice quality and vocal function. Once again it appears that patients before radiotherapy experienced decreased vocal performance, which improved for patients six months to seven years

after radiation but remained worse than vocal performance as reported by control speakers. Also, it appears again that diagnostic stripping instead of biopsying the vocal folds and continuing smoking after treatment have an adverse effect on vocal performance following radiotherapy.

The conclusion of this thesis (Chapter 9) is that voice characteristics remain worse for almost half of the patients six months to seven years after radiotherapy compared to control speakers. Carefully balancing the advantage and disadvantage of stripping the vocal fold for initial diagnosis and emphasising the negative effect of continuing smoking is thereby of interest. Furthermore, it appears that because of the multidimensional character of voice, an analysis protocol should comprise multiple voice measures. Based on the findings in this thesis, this protocol should comprise at least perceptual ratings of voice quality by trained raters on running speech, preferably complemented with acoustical measures, evaluations of stroboscopic video-recordings of vocal function, and self-ratings of vocal performance. Although more research is needed on reliability, validity, and feasibility of (other) voice analysis methods, this concept protocol is useful in clinical studies on the evaluation of treatment for patients diagnosed with early glottic cancer.

FUNCTIONAL PHONOLOGY: FORMALIZING THE INTERACTIONS BETWEEN ARTICULATORY AND PERCEPTUAL DRIVES

author: Paul Boersma
promotor: Louis C.W. Pols
date of defence: September 14, 1998

Summary

In this book, I showed that descriptions of the phenomena of phonology would be well served if they were based on accounts of articulatory and perceptual needs of speakers and listeners. For instance, the articulatory gain in pronouncing an underlying |n+k| as [ŋk] is the loss of a tongue-tip gesture. Languages that perform this assimilation apparently weigh this articulatory gain higher than the perceptual loss of the coronal place cues. This perceptual loss causes the listener to have more trouble in reconstructing the perceived /ŋ/ as an underlying |n|. This functionalist account is supported by the markedness relations that it predicts: the ranking of the faithfulness (anti-perceptual-loss) constraints depends on the perceptual distance between the underlying specification (/n/) and the perceptual result (/ŋ/) and on the commonness of the feature values (coronal is more common than dorsal), leading to more or less fixed local rankings as

"do not replace /t/ with /k/" >> "do not replace /n/ with /ŋ/"

and

"do not replace /ŋ/ with /n/" >> "do not replace /n/ with /ŋ/"

where the " >> " symbol means "is ranked higher than" or "is more important than". The first of these two rankings is universal because plosives have better place cues than nasals, and the second is valid in those languages where coronals are more common than dorsals (ch. 9). These universal rankings lead again to near-universals (ch.11) like "if plosives assimilate, so do nasals (at the same place of articulation)" and "if dorsals assimilate, so do coronals (in languages where coronals are more common than dorsals)".

The idea of constraint ranking is taken from Optimality Theory, which originated in the generative tradition (Prince & Smolensky 1993). The interesting thing of the optimality-theoretic approach to functional principles, is that phonetic explanations can be expressed directly in the production grammar as interactions of gestural and faithfulness constraints. This move makes phonetic explanation relevant for the phonological description of how a speaker generates the surface form from the underlying form. I have shown (chs. 13, 17, 18, 19) that this is not only a nice idea, but actually describes many phonological processes more adequately than the generative (nativist) approach does, at least those processes that have traditionally been handled with accounts that use the hybrid features of autosegmental phonology, underspecification theory, and feature geometry.

The model of a production grammar in functional phonology (ch. 6) starts with a *perceptual specification*, which is an underlying form cast in perceptual features and their combinations. For each perceptual specification, a number of *candidate articulations* are evaluated for their articulatory effort and for the faithfulness of their *perceptual results* to the specification. This evaluation is performed by a grammar of many strictly ranked articulatory constraints (ch. 7) and faithfulness constraints (ch. 9), and the best candidate is chosen as the one that will be actually spoken.

There is also a perception grammar, which is a system that categorizes the acoustic input to the listener's ear into language-specific perceptual classes (ch. 8). The listener uses the perception grammar as an input to her speech-recognition system, and the speaker uses the perception grammar to monitor her own speech: in the production grammar, a faithfulness constraint is violated if the output, **as perceived by the speaker**, is different from the specification.

In the language-learning child (ch. 14), the production and perception grammars are empty: they contain no constraints at all. As soon as the child acquires the categorization of acoustic events into communicatively relevant classes, the perception grammar comes into being, and as soon as the child decides that she wants to use the acquired categories to convey semantic and pragmatic content, faithfulness constraints arise in the production grammar. As soon as the child has learned (by play) how to produce the required sounds, constraints against the relevant articulations enter the production grammar. These constraints lower as the child becomes more proficient (by play and imitation), thus leading to more faithful utterances. A general gradual learning algorithm hypothesizes that the child will change her constraint rankings (by a small amount) if her own utterance, **as perceived by herself**, is different from the adult utterance, **as perceived by the child** (the bold phrases on this page stress the prominent role for perception in a functional theory of phonology, as opposed to theories that maintain hybrid phonological representations). This learning algorithm, by the way, is capable of learning *stochastic grammars*, i.e. the child will learn to show the same degree of variation and optionality as she hears in her language environment (ch. 15).

The original aim of this book was to propose a model for inventories of consonants, based on functional principles of human communication, like minimization of articulatory effort and minimization of perceptual confusion. The symmetry that phonologists see in these inventories follows from the finiteness of the number of perceptual categories and the finiteness of the number of acquired articulatory gestures. The gaps that phoneticians see in these inventories follow from asymmetries in the context dependence of articulatory effort and perceptual contrast. This functional approach to inventories (ch. 16) and phonological phenomena in general marries the linguist's preference for description with the speech scientist's preference for explanation, in a way that, I hope, will eventually appeal to both convictions.

SPEECH VARIABILITY AND EMOTION: PRODUCTION AND PERCEPTION

author: Sylvie Mozziconacci
promotores: Adrian J.M. Houtsma & Louis C.W. Pols
copromotor: Dik J. Hermes
date of defence: November 20, 1998

Summary

Experiences in every-day life illustrate that the contents of spoken communication are not restricted to *what is said*, but also involve *how it is said*. A huge number of variations occur in speech, so that saying a sentence twice does never result in exactly the same acoustic realization. This might lead a listener to interpret the two utterances as two different messages. Speakers exploit this freedom to vary speech components in order to express themselves, and listeners take this variation into account when decoding the spoken message. Today's speech-synthesis systems do not compare with humans, even remotely, when it comes to exploiting prosodic variation. As a consequence, today's synthetic speech, despite the fact that it is considered reasonably intelligible, is also perceived as dull. It sounds rather unnatural and uninvolved. Modeling variability in synthetic speech is expected to enhance its quality and, therefore, to increase its potential use. The scale of variation involved in speech produced in emotional states, is wide. Acquiring knowledge concerning these variations is expected to make it possible to model speech variation associated with emotion, as well as to model more moderate variation that is not so much associated with emotional involvement, but rather with enhancing naturalness in neutral utterances.

In the present study, the variation of the prosodic elements: pitch level, pitch range, intonation pattern, and speech rate was investigated in the vocal expression of emotion. These parameters are considered to have a major contribution in conveying emotions. In order to be able to use the results of the present study in speech synthesis, it is of relevance not only to describe the speech variation qualitatively, but also to quantify it. Since utterances conveying neutrality are the usual output of speech-synthesis systems, it is also convenient to express variation in parameter values in terms of deviation from neutrality. In order to model only the speech variability as far as it is relevant to communication, the present investigations do not only include *production* studies, but also *perception* studies. An experimental approach is used, in which analyses of natural speech variation are carried out and perceptual tests involving synthetic or re-synthezised speech are performed, in order to test the relevance of the data found. Furthermore, the consideration of these variations in the framework of models commonly used in speech studies, allows the validity of these models to be tested.

In Chapter I, the problems at hand are described. The framework, in which studies concerned with the expression of emotion in speech are carried out, is depicted, approaches are discussed, and the approach adopted for the present study is presented. Finally, an outline of the investigation is given.

Chapter II deals with the selection of the speech material for use in the present study. The selection of 315 utterances (3 speakers × 5 sentences × 7 emotions × 3

trials) was based on appropriate emotion identifiability. A representative subset of these, consisting of 14 utterances (1 speaker × 2 sentences × 7 emotions × 1 trial), was intended for use in the preliminary analyses of Chapter II. The seven emotions: 'neutrality', 'joy', 'boredom', 'anger', 'sadness', 'fear', and 'indignation', were involved in the present investigation. The identification of these seven emotions in the original speech was tested in a perception test. The results form a useful basis for comparison with the results of later experiments. Next, the adequacy of the semantic content of the five sentences for use in this study was tested and confirmed. An analysis of the subset of fourteen utterances was then carried out at utterance level, by means of measurements of pitch level, pitch range, and speech rate. Additionally, these fourteen utterances were individually labeled in terms of intonation patterns, according to the Dutch grammar of intonation by 't Hart, Collier and Cohen (1990). A series of experiments was conducted in which pitch level, pitch range, and speech rate were systematically varied, per emotion, around the values found for these parameters in the original speech. The variation in intonation patterns was controlled by providing each test utterance with the same intonation pattern as in the original utterance of the corresponding emotion. Perception experiments were carried out, in which subjects ranked the utterances they found best for the expression of a specific emotion. On the basis of the results, optimal values for pitch level, pitch range, and speech rate were derived for the generation of emotional speech from a neutral utterance. These values were then perceptually tested, in experiments in which subjects labeled utterances with the name of one of the seven emotions. The first series of experiments involved resynthesized speech, while the last experiment involved rule-based synthetic speech. Applying the values that were found optimal, onto synthetic speech, lead to a good identification of the emotions, namely 63% correct identification. Although some emotions were less successfully identified than others, general results were quite encouraging. Results showed that pitch and speech rate are powerful cues for conveying emotion in speech.

In Chapter III, an extensive study was conducted, concerned with F_0 fluctuations produced in the expression of emotion, and with the relevance of perceived pitch variations for the identification of emotion in speech. Pitch level and pitch range were estimated on the basis of measurements of mean F_0 and its standard deviation in the 315 utterances in the database. It was shown that, after speaker normalization, the values found in natural utterances produced by the three speakers eliciting the seven emotions, closely matched the optimal values obtained in the perception tests of Chapter II. The course of pitch in all individual utterances was described in terms of the model of intonation by 't Hart, Collier and Cohen (1990), describing a pitch curve as a combination of a slowly decreasing component (the declination line) and relatively fast pitch movements, superimposed on this baseline. In this model, the end point of the declination line represents the pitch level, while the excursion size of the pitch movements represents the pitch range. In principle, this excursion size of the pitch movements is considered to be constant throughout the utterance, so that pitch curves could also be described with a lower declination line, or baseline, and an upper declination line, or topline, between which the pitch movements are realized. The overall excursion size of the pitch movements then equals the distance between the lower and the upper declination line. In Chapter III, the relationship was discussed between two ways of estimating pitch level and pitch range. One estimation was model-based, involving the end point of the baseline and the difference between baseline and topline, respectively. The other estimation, more

strictly data oriented, was based on the average of F_0 in the utterances and the standard deviation of F_0, respectively. Furthermore, pitch level and pitch range can only be defined as properties over the whole utterance. In naturally produced pitch curves, many details can be distinguished which cannot be captured in such a model of intonation. In order to study the fluctuations of F_0 occurring within utterances, F_0 was measured at a number of fixed points in the utterances. Measurements were carried out in the first voiced part of the utterance, in the vowel of the first accent peak, in a vowel after the initial accent peak, in a vowel before the final accent peak, in the vowel of the last peak, and in the last voiced segments of the utterance. It appeared that utterances produced while conveying different emotions could vary considerably with regards to relative peak heights and the extent of final lowering. For instance, the F_0 measurements concerning the last accent peak often yielded a higher value than the measurements concerning the first peak, which cannot be accounted for on the basis of declination only. Especially for some emotions, the final measurement of F_0 yielded a lower value than could be expected on the basis of preceding measurement of F_0 that are expected to be representative of the baseline. In a perception study, the relevance of these differences was put to the test. Although some effects appeared to be significant, e.g., modeling final lowering appeared to increase the number of responses of the subjects indicating indignation, the effects found were relatively small.

The 315 utterances selected as speech material were labeled in terms of intonation patterns, and the distribution of the patterns of pitch movements over the various emotions was investigated per speaker. The results are presented in Chapter IV. It appeared that the patterns were not equally distributed over all seven emotions. The '1&A' pattern, a prominence-lending rise-fall, was the most often used pattern; it was regularly produced in all seven emotions. Therefore, the hypothesis emerged that this '1&A' pattern would be a good candidate to apply to all emotions, so that no variability is introduced by the realization of different intonation patterns. From the production study, however, it also appeared that many utterances were produced with other intonation patterns, and some intonation patterns seemed to be more characteristic for some emotions than for others. In particular, it was noticed that the patterns '12' (a rise followed by a very late rise) and '3C' (a late rise and a very late fall), were never used in final position in utterances expressing neutrality. A second hypothesis, therefore, emerged concerning the question of whether the two patterns '12' and '3C' could signal emotion in speech. A perception experiment was carried out, investigating the perceptual relevance of intonation patterns for identifying emotions in speech. This test provided converging evidence on the contribution of specific patterns in the perception of some of the emotions studied. Some intonation patterns introduced a perceptual bias towards a specific emotion. Finally, clusters of intonation patterns were derived from the results of the perception experiment. The last part of the pattern appeared to be of particular relevance. The clustering reflected the perceptual distinctions among intonation patterns.

In Chapter V, temporal variations conveying emotion in speech were investigated. First, an analysis of speech rate was performed at utterance level. Global measurements of overall sentence duration and its standard deviation were carried out on the 315 utterances selected as speech material. Averages per emotion were calculated for each speaker. It was investigated whether a linear approach, simply consisting of stretching or shrinking the whole utterance linearly, i.e., manipulating the overall speech rate, is sufficient for expressing emotion in speech, or whether a more detailed approach would be necessary. To this end, an

analysis was performed below utterance level. Measurements of relative duration of accented and unaccented speech segments (syllables or groups of syllbales) were made, in order to acquire some insight into the internal temporal structure of emotional utterances. Although differences are small and the analysis of production data did not provide conclusive evidence of the systematic use of variation in the internal temporal structure of utterances in speech conveying an emotion, some of the detailed information could not be described with a linear-stretch model. The perceptual relevance of separately stretching or shrinking speech segments within utterances was then questioned. The deviation from a linear model could either specifically be due to the expression of emotion, or simply due to the modification of overall speech rate and, therefore, be only indirectly related to the expression of emotion (i.e., only because emotion is conveyed with changes in overall speech rate). In order to obtain the reference required for deciding which interpretation is correct, the same measurements of relative duration of accented and unaccented speech segments were made in neutral speech, spoken at different overall speech rates, by one of the male speakers. The results of the measurements in emotional and in neutral speech were compared. The temporal structure appeared to change non-linearly and to vary with some of the emotions. An experiment was carried out in order to test the perceptual relevance of these variations. Speech manipulations were carried out in order to generate emotional speech, either by simply stretching or shrinking the whole utterance linearly, or by proportionally varying the duration of accented and unaccented speech segments. Values for relative durations tested in the experiment were inspired from the production data. The differences in relative duration of accented and unaccented speech segments that are *associated with speech rate,* appeared not to be perceptually relevant. On the other hand, the differences in relative duration of accented and unaccented speech segments that are *associated with the expression of emotion,* appeared to be perceptually very relevant for the expression of neutrality and indignation.

Finally, in Chapter VI, the limited research area of the present investigation is once again justified and the results of the study are summarized. It is concluded that an interaction of some prosodic cues permits the vocal expression of emotion, and that most emotions can be conveyed in synthetic speech by controlling the parameters studied here. For some emotions, however, this is less successful. For these emotions, other cues, such as voice quality, loudness or other properties of intonation, may be essential. The results that were found to be specific to the expression of emotion in speech are given as a series of rules for generating speech in each of the emotions studied. These rules are summarized in the table presented above , in which optimal values are mentioned for each emotion. A specification is also given of which patterns are preferred or should be avoided in the modeling of the emotions, and whether or not a modeling of final lowering and relative height of the peaks is expected to be relevant.

Additionally, general results concerning the suitability of models for handling the extreme variations occurring in emotional speech were summarized. The thesis is concluded by some suggestions of lines for future research concerned with the expression of emotion in speech.

Relationships established between emotions and parameters, based on the production and/or on the perception studies

Parameters	Emotions						
	neutrality	joy	boredom	anger	sadness	fear	indignation
pitch level	65 Hz	155 Hz	65 Hz	110 Hz	102 Hz	200 Hz	170 Hz
pitch range	5 s.t.	10 s.t.	4 s.t.	10 s.t.	7 s.t.	8 s.t.	10 s.t.
final lowering	-	-	no	-	yes	yes	yes
relative peak height	-	-	-	-	yes	yes	-
pattern(s) to prefer in final position	1&A	1&A and 5&A	3C	5&A, A and EA	3C	12 and 3C	especially but also 3
pattern(s) to avoid in final position	12 and 3C	A, EA, and 12	5&A and 12	1&A and 3C	5&A	A and EA	1&A
duration relative to neutrality	100%	83%	150%	79%	129%	89%	117%
durational proportion acc./unacc. segments	no deviation from linearity	-	-	-	-	-	stretch acc. segments 40% more than unacc.