Some phonetic experiments on: Double stress and rhythmic variation in R.P. English

Heuven, V.J.J.P. van

Some phonetic experiments on:

Double stress and rhythmic variation in R.P. English

Vincent J. van Heuven

Utrecht University
University of Edinburgh
1974

Preface

The present thesis is an account of an experimental study on double stress in English. The original idea was to approach this phenomenon from three angles, the phonetic angle: “Does it exist, and if so, what does it look like?”, the linguistic angle, rather narrowly interpreted as “Can it be described within a transformational generative framework?”, and finally the applied angle “How to teach double stress to Dutch learners of English?”

Clearly, the latter two aspects are only relevant if it can be demonstrated that double stress does exist, and in this respect the phonetic aspect is primary.

Last year I wrote a paper covering the linguistic aspect, in which I suggested two competing sets of rules to be added to those given in Chomsky and Halle (1968), one of which would apply in case double stress does exist, the other if the assumption should turn out to be false.

The applied aspects have not been dealt with until this moment; the phonetic side, however, is the subject of the investigation reported on in this paper.

The topics were suggested to me by Prof. A. Cohen of Utrecht University; the research was carried out at the University of Edinburgh, Scotland, where I stayed during the academic year 1972/73. The report was written in Utrecht. Supervisors in Utrecht were A. Cohen and M. van den Broecke, while part of their responsibility was taken over by J. Antony and L. Iles of Edinburgh University.

Contrary to the requirements for doctoral theses at the English Dept. of Utrecht University, this paper is probably not fully understandable to uninitiated readers, for which I apologize. I hasten to point out, however, that the general idea should not be difficult to grasp, especially when the various references to some introductory works on the subject are followed up.

Let me finally thank a number of people at the Universities of Utrecht and Edinburgh who have advised, taught, assisted or stimulated me in the course of this work (the order is alphabetical): J. Antony, M. van den Broecke, H. Cirkel, A. Cohen, D. Cruickshank, L. lles, I. MacVey-Gow, R. Motherwell (!), S. Stephens, J. Laver, Mrs. E. Uldall, and Mrs. R. Clark.

I have purposely avoided specifying what each of these people have contributed to this paper so as not to create the impression that my own part was to sit back and watch other people do the work for me. Special thanks are due to the Students to England Committee and the Dutch Ministry of Education, who made my stay in Edinburgh possible, and to my wife Petra, who interrupted her studies to go with me.

Soesterberg, Netherlands, February, 1974

Contents

Chapter one

Introduction

1.1 Stress

1.1.1 Definition

By stress I will mean the relative amount of physiological effort that has gone into the production of a syllable. The effort may be applied in the pulmonary, phonatory, and articulatory stages of speech production, but it is not known if there is an order of importance among these three (Ladefoged 1971: 83, Öhman 1967: 47, Netsell 1970).

On the auditory level, stress is the subjective impression on the part of the listener of the relative amount of effort the speaker uses in the production of a syllable.

There is no acoustic factor or complex of factors that can be closely associated with perceived stress (Lehiste 1970: 110). Stress tends to coincide with higher values on the fundamental frequency, intensity and duration parameters.

It is an open question whether stress equals prominence. I have allowed for the possibility that intuitive corrections for inherent sonority are applied by listeners, which would separate stress and prominence (Lehiste and Peterson 1960, Lehiste 1970: 118).

1.1.2 Stress levels

Defined in this way, the number of stress gradations is practically unlimited. It is customary, however, to postulate a number of stress levels, sufficient to make an adequate description of a stress system possible.

English is said to have several stress levels, the exact number ranging from two (stressed and unstressed) to indefinitely many. In the majority of the handbooks three or four levels are distinguished: strong - medium - (weak) - unstressed. The systematic phonetic stress representations in generative phonology, which use indefinitely many levels, are based on three-level transcriptions (Kenyon and Knott 1944).

1.1.3 Stress patterns

By stress pattern we shall mean the succession of various stress levels within a word.

In this investigation I have restricted myself to two-syllable words mainly for statistical reasons (cf. van Heuven 1972).

I shall adopt a rather conventional notational system for stress levels and patterns, where strong (or primary) stress is represented by (1), medium (or secondary) stress by (2), and weak stress by (3), i.e., where lower degrees of stress are symbolized by higher integers. Stress patterns of two-syllable words will be represented by hyphenated pairs of integers: (1-2) would stand for a two-syllable word with strong stress on the first syllable, and medium stress on the second.

1.2 Stress patterns in English

On the basis of three stress levels and two-syllable words, six patterns can be produced. The (3-3) and (2-2) patterns have to be omitted from further discussion, as these are nowhere to be found in the literature on the subject. As a matter of fact, it seems to be a tacit assumption that there must always be at least one primary stress in a word, which, of course, simply rules out combinations of the above type.

It is generally agreed upon that three of the remaining combinations are regularly used in English, viz. (1-3), (1-2), and (3-1) patterns (for examples cf. Gimson 1962: 228).

Opinions diverge on the (1-1) and (2-1) patterns. According to the − predominantly − British tradition there are (2-1) words, although relatively few (for an exhaustive list cf. Kingdon 1958a: 196), as well as (1-1) words, which then is a quite frequently used pattern.

According to the other tradition, which has most of its adherents among American phoneticians and phonologists, there is only one pattern at stake, viz. the (2-1) pattern; i.e. the (1-1) pattern does not exist.

These conflicting views cannot be reduced to stress differences in the British and American dialects of English, as I have argued elsewhere. For a more detailed discussion and extensive bibliography I refer to van Heuven (1973).

1.3 Double stress

The controversial (1-1) pattern is known as double stress, level stress, even stress, and equal stress. I shall use these terms indiscriminately.

Scholars who believe that double stress exists, seem to imply that this pattern is exceptional. Thus, double stressed words are said to constitute “an unexpectedly large proportion of the English vocabulary” on one occasion (Kingdon 1958a: 15), and to be “relatively rare in English, although the absolute number of cases in Jones’ Dictionary is not very small” on another (Vanvik 1964: 66).

Also, it is often intimated that double stress is an exclusively English phenomenon. Many of the handbooks include a section called “advice to foreign learners”, and instructions are given how to pronounce two equal stresses. Double stress appears to occur in at least two other related languages, viz. German (von Essen 1966) and Dutch (Kruisinga 1927).

1.4 The rhythmic principle

Double stressed words, and only these, are allegedly subject to what has been called the rhythmic principle.

When the rhythmic principle is defined on (1-1) words, it asserts that:

(1) when a (1-1) word is preceded by another strong stress, without any intervening weak stresses, its first strong stress is lowered to a medium stress, giving a (2-1) pattern;

(2) when such a word is immediately followed by a strong stress, the second of the two strong stresses becomes medium stress, giving a (1-2) pattern;

(3) when both preceded and followed by strong stresses, either case (1) or case (2) applies, depending on which of the two words has a closer grammatical relation with the double stressed word (Kingdon 1958a: 165, van Heuven 1973: 29);

(4) in all other contexts, i.e. when surrounded by unstressed syllables, or when spoken in isolation, double stressed words are actually realized as (1-1) patterns. Vanvik, however, (1962: 67) claims that the (1-1) realization has to be excluded in citation forms as well.

The rhythmic principle applies on a more limited scale when the existence of the (1-1) category is denied from the beginning: here the change from (1-1) to (2-1) is impossible. Within the “American” tradition its only effect is to invert a (2-1) pattern to (1-2) when a strong stress immediately follows; in all other contexts the original (2-1) pattern is preserved (Kurath 1963: 142).

1.5 Aims of this investigation

The basic aim of this study is to shed some light on these partially conflicting allegations. In its crudest formulation, what I want is to find out if there is such a (1-1) stress pattern, and if so, what it looks like.

More generally, I shall try to find out experimentally if there are five different two-syllable-word stress patterns, where (1-1) and (2-1) are distinct categories, or only four, where these two coincide.

1.6 Basic considerations

Obviously, it will not suffice simply to consider the (1-1) category in isolation, and see if the two stresses are exactly balanced, though this in itself is an interesting question. Should it turn out that the two stresses are not exactly equal, there is still the weaker interpretation that the distribution of stress over the two syllables approaches the equilibrium more optimally than in any of the other patterns. Such an interpretation is, in fact, intimated by Kenyon and Knott (1944: xxi) when they define double stress as the occurrence of two equal or nearly equal stresses in one word. So, a second, and in view of the above considerations, more realistic approach is to concentrate on the difference between the (1-1) and (2-1) patterns.

It is a fortunate circumstance that we can now appeal to the rhythmic principle. In the ‘British’ tradition it generates (2-1), (1-1) and (1-2) stress patterns on the same lexical material, if the word concerned is of the double stressed type. Should we never obtain any evidence to the effect that there are systematic differences between the (1-1) and (2-1) realizations of such words, we may safely assume that a description of the English word stress system in terms of four categories without the double stress pattern is preferable.

Chapter two: Some assumptions;

Orientation towards the literature

2.0 Introduction

A number of preliminary decisions had to be taken before I could begin experimenting. Because these are of a fundamental, rather than a merely practical or instrumental, nature, I prefer to discuss them under a separate heading, instead of dealing with them as they come up in the individual reports on the various experiments.

2.1 Dialects

Although all the claims in chapter I seem to pertain to every dialect of English (Fuhrken 1934: 85), I have limited the scope of the investigation to Standard British (R.P.) English, for practical reasons only. A good deal of phonetic research on English is based on this variety, so that it would be unwise not to follow this procedure, unless for contrastive purposes, which motive was absent from this set-up.

2.2 Sorts of evidence

2.2.0 Introduction

In § l.1 I have given definitions of stress in terms of its production, acoustic manifestation, and perception.

In principle, we can look for evidence relevant to questions concerning stress in each of these three areas. In practice, however, I have deliberately avoided this line of action. In the next few sub-sections I shall briefly state my reasons for doing so.

2.2.1 Evidence from speech production

My reasons for not looking for evidence in this area are twofold: Firstly, the experimental techniques that one would have to apply here, such as electromyography, measuring sub-glottal (tracheal or oesophageal) air pressure (cf. Lehiste 1970: 108), are of a highly sophisticated type, and beyond my reach at the time that I started on this investigation.

Secondly, it appears that data obtained from these techniques can only in a very rough way be correlated with stress; the distinction between strongly stressed and unstressed syllables can be made, but it is as yet impossible to set up a rank-order of stress levels on the basis of these data. This, however, is precisely what I am after, and for this reason I decided not to consider physiological evidence any further.

2.2.2 Acoustic evidence

As I have said in § 1.1, it is still not known exactly in what way stress production (i.e. the application of extra effort, effectuating an increase in subglottal air pressure) is manifested acoustically, nor what acoustic factor, or factors, are responsible for the perception of stress. In particular, suggestions and proposals concerning the trade-off relationships among the various parameters have been unsatisfactory up to this very moment.

In spite of these considerations, however, I have decided to use at least some evidence of this kind in my investigation.

First of all, the techniques involved are rather simple, and the results can be stated in clear-cut physical measures which gives a firm basis for further research.

Secondly, the variability on these acoustic parameters is such that we may hope, in principle at least, to obtain a more refined classification among syllables than the stress/unstress distinction. Naturally the results of these experiments will have to be treated with necessary caution.

The final, and most important, reason for including an acoustically oriented experiment was the fact that it was to serve as a necessary preliminary for further perceptual experiments. This aspect is discussed in more detail in §§ 4.5.3 and 5.0.

2.2.3 Perceptual evidence

The ultimate decision whether a syllable is stressed or not (or somewhere in between these extremes) resides with the listener. This aspect is primary because, before we start investigating physiological and acoustic properties of stressed syllables, we have to know that they are stressed in the first place.

In view of the difficulties researchers have experienced in defining stress, and in stating its productive and acoustic correlates, it is remarkable how easily and consistently native speakers are able to tell stressed from unstressed syllables, when confronted with speech samples of their own language.

Precisely because of the theoretical priority and the technical feasibility of perception tests I have decided to concentrate my attempts at solving the question of double stress on getting evidence from perceptual data.

2.2.4 The importance of synthetic speech

Clearly, it would be unwise to use samples of naturally produced speech for such perception tests. If a subject considers a particular syllable stressed, this may be due to any of a number of factors. For instance, he may perceive stress because there is a momentary rise in the fundamental frequency, or alternatively, he may find it unstressed because the syllable is shorter than normal, in spite of the increased fundamental frequency. In fact, the number of variations in speech signals is unlimited and we do not know which variations govern stress perception. There may very well be relevant properties of the acoustic signal we have not yet bothered to think about.

As long as we do not know exactly which parameters are responsible for stress perception, and what their trade-off is, using natural speech will always be hazardous.

It has therefore become a standard procedure to use synthetic speech for phonetic research of this kind. Here we know, and decide for ourselves, exactly what our speech samples will look like. We can avoid possible trade-off relationships by varying only one relevant parameter at the time, or choosing fixed relations among the parameters and, finally, we can vary a particular parameter with infinitely more precision than a human voice could ever do this.

For these reasons I have based my crucial experiment on synthetic stimuli. I have, however, also included reports on perception tests with natural speech; my motivation for carrying these out was curiosity rather than aspiration to experimental validity, and therefore they are of a non-decisive, in fact, marginally relevant nature (see also chapter III).

2.3 Stress as a binary vs. multi-valued distinction

2.3.0 Introduction

As I have tried to make clear in §§ 1.2 and 1.3, the problem this paper tries to come to grips with, is a matter of stress levels and patterns, rather than the simple distinction between stressed and unstressed.

The vast majority of the literature has concerned itself with establishing the physiological and acoustic correlates of stress as opposed to non-stress. If we want to compare stress patterns, that is to say, a succession of stress levels within one word, we will obviously need a more refined classification.

The evidence that a multi-valued stress distinction is at all possible is rather meagre, mainly, I take it, because thin aspect has not received much attention so far.

In the following subsections I will briefly review what has been reached in each of the three basic areas of research.

2.3.1 Review of experiments

2.3.1.1 Physiological

I know of no serious attempts to establish a hierarchy of stress on the basis of, say, electromyographic data. As stated in § 2.2.1, such an analysis has not yet proceeded beyond a two-way classification. Moreover, since physiological data will not be taken into account in this investigation, we will not go into this matter in any detail.

2.3.1.2 Acoustic

In the literature I have surveyed in the course of this investigation I have come across two experts who were concerned with establishing acoustic correlates of more than two stress levels.

Lieberman (1967: 150) reports on an experiment in which he tried to find evidence for the existence of an intermediate stress level. He claims that the relevant cues were pause phenomena, parameters which I have not included in my experiments.

McAlister (1971) conducted experts to find acoustic differences of a gradual nature among the various stress levels predicted by the transformational cycle (cf. Chomsky and Halle 1968, Halle and Keyser 1971). He claims that hierarchical ordering of stresses can be based on acoustic parameters, at least to a limited extent.

2.3.1.3 Perceptual

Experiments involving natural speech tend to support the view that speakers of the language concerned are able to make a systematic distinction among a number of stress levels (Kost, Zinkstok, and Zonneveld 1972). Lieberman, however, suggests that this ability resides with the listeners' knowledge of the language, and that it is not governed in any significant way by what is acoustically present in the signals. When the lexical information was eliminated from the utterances by vocoder synthesis techniques, no more than two stress levels (stressed and non-stressed) could be detected by linguistic experts (Lieberman 1965).

2.3.2 Implications

On the one hand, the assumption that stress can be conceived of as a multi-valued scale, can be met with reasonable optimism; on the other hand, it seems to me that the most important justification of this assumption will have to be given by this investigation itself. I believe, however, that the results of my experiments show that the assumption is reasonable.

What techniques have been used to elicit such refined distinctions among stresses will be dealt with as we come to them in the reports on the various experiments.

Chapter three

Organization of the rest of this thesis

For the sake of clarity I have divided this paper into two halves, viz. reports on central experiments, and reports on peripheral experiments.

Part 1 comprises a series of loosely interrelated experiments, which, when taken as a whole, have a direct relation to the question whether or not double stress exists. Exactly how they are interrelated will be explained in the introductory section to the individual experiments. This series is self-contained, and the three other peripheral experiments could very well have been left out, as they are only marginally relevant: at the most they add some extra support to decisions taken in the central experiments.

I have decided to include them all the same for the following reasons: This paper is not just a report on an investigation; it is also a survey of what I have done during my stay at Edinburgh University. Having spent about four months' time on the peripheral experiments, I felt that leaving them out would be an incorrect reflection of my activities there. Secondly, on the occasion of an informal lecture on my work on the peripheral experiments many people appeared to be interested, and asked if they could get a written version of the final report on this work.

It should be pointed out that the reports in this paper have not been given in their chronological order. There was a time lag of four months between Experiment I "analysis" and Experiment II "synthesis". The peripheral experiments were designed and carried out in this period.

Chapter four

Central experiments: Analysis

4.0 Introduction

This experiment was devised to give us a rough indication as to what the various stress patterns involved in this investigation look like. As such it was a necessary preliminary to my main experiment.

It is a rather common procedure to base one's perceptual tests on the findings of a preceding analytic experiment (Fry 1955, Lehto 1969).We have followed this procedure here.

4.1 Stimuli

The words *unknown*, *eighteen*, and *mince pie*, supposedly representing the class of double stressed words, *window* and *footprint*, *absurd* and *machine*, representing falling and rising patterns, respectively, were fitted in five phonologically different environments. These words, the categories they belong to, and the phonological environments are given matrix-wise in Table 1.

The choice of the double stressed words was based on the criterion that they be typical recurrent examples of double stress in the majority of the handbooks. The falling stresses are of the (1-2) type, which is closer to double stress (1-1) than any other pattern. The words with rising patterns are usually transcribed with a (3-1) contour. Admittedly, there are some instances of words with (2-1) patterns (for an exhaustive list see: Kingdon 1958: 196) but it proved to be impossible to fit these in the intended phono-syntactic environments. The (3-1) words were therefore chosen to represent the rising pattern closest to double stress.

The five phonological environments represent instances of

- Preceding strong stress : 1_0
- Following strong stress : 0_1
- Both preceding and following strong stress : 1_1
- Neither preceding nor following strong stress : 0_0
- Citation form or lexical pronunciation : #_#

The 7 × 5 sentences (and in the case of citation forms: words) were typed out on individual cards and these were ordered in such a way that instances of the same word or phono-syntactic environment never clustered. This was done to conceal the intention behind the experiment from the subjects as much as possible. An exception to this rule were the words in citation form, which had to be ordered at the end of the series.

Table 1: Target words, stress patterns, phono-syntactic environment and order of presentation.

Pattern	Environment	Order	Sentence
(1-1)	0_0	1	Things like that are **unknown** in this country
(2-1)	1_0	23	Things like that are quite **unknown** in this country
(1-2)	0_1	20	Things like that are **unknown** objects in this country
(2-1)	1_1	26	Things like that are quite **unknown** objects in this country
(1-1)	#_#	31	**Unknown**
(1-1)	0_0	21	She was **eighteen** at the time
(2-1)	1_0	15	She was just **eighteen** at the time
(1-2)	0_1	12	There were **eighteen** girls at the party
(2-1)	1_1	18	There were just **eighteen** girls at the party
(1-1)	#_#	34	**Eighteen**
(1-1)	0_0	17	We are having **mince pie** for dinner
(2-1)	1_0	11	We'll have a hot **mince pie** for dinner
(1-2)	0_1	8	I ate the **mince pie** hot at dinner yesterday
(2-1)	1_1	14	I'll have a hot **mince pie** first thing in the morning
(1-1)	#_#	35	**Mince pie**
(1-2)	0_0	25	He jumped from the **window** on the first floor
(1-2)	1_0	19	He jumped from the right **window** on the first floor
(1-2)	0_1	16	He jumped from the **window** just in time
(1-2)	1_1	22	He jumped from right **window** just in time
(1-2)	#_#	30	**Window**
(1-2)	0_0	9	I looked at the **footprint** in the garden
(1-2)	1_0	3	I saw a clear **footprint** in the garden
(1-2)	0_1	28	There was a **footprint** right on the spot
(1-2)	1_1	6	There was a clear **footprint** right on the spot
(1-2)	#_#	32	**Footprint**
(3-1)	0_0	13	It is rather **absurd** to say it
(3-1)	1_0	7	It is quite **absurd** to say it
(3-1)	0_1	4	It is an **absurd** thing to say
(3-1)	1_1	10	It is a quite **absurd** thing to say (3-1)
(3-1)	#_#	33	**Absurd**
(3-1)	0_0	5	He'll get the **machine** in the morning
(3-1)	1_0	27	He'll get the new **machine** in the morning
(3-1)	0_1	24	He'll get the **machine** back in the morning
(3-1)	1_1	2	He'll get the new **machine** back in the morning
(3-1)	#_#	29	**Machine**

4.2 Subjects

Subjects were five male native speakers of English (ages: 20, 22, 23, 24, and 38) chosen on the criteria of availability and their being speakers of (at least a reasonable approximation to) R.P.-English. They were four students and one lecturer at Edinburgh University, and none of them was linguistically naive. They cooperated on a voluntary basis, and were not paid.

4.3 Procedure

The subjects were instructed to read out the sentences on the cards one by one. They could take one good look at each sentence immediately before reading it out. They were told not to stammer or hesitate once they had started reading out a particular sentence. In case a sentence came out unsatisfactorily, it had to be repeated at once. No other instructions were included.

Microphone and laryngograph (glottograph) outputs were simultaneously recorded on separate channels of a tape recorder. The laryngograph signal was used to control a pulse generator, and it was this signal that was in fact recorded. The laryngograph was used to arrive at more reliable and accurate measurements of the fundamental frequency. A more detailed description of the laryngograph can be found in Fourcin and Aberton (1971: 172-182).

4.4 Analysis

4.4.1 Instrumental analysis

The recordings were edited in order to compress the quantity of data to be analysed. The laryngograph signal was then fed into a Frøkjær-Jensen Trans Pitch meter, while the microphone output was fed into a combined intensity meter/oscillograph manufactured by the same company as above. The output of these apparatus was simultaneously recorded on a four-channel mingograph at 10 cm/sec; for a description of these instruments see Fant (1958), mingograms are included in appendix I. There the bottom trace is a time calibration, where each complete oscillation corresponds to 50 msec. The lower middle trace is an oscillogram of the microphone signal, which was included to facilitate segmentation. The upper trace is an intensity graph of the microphone signal; calibrations are given in Figure 1, integration time 20 msec. The upper middle trace, finally, is the laryngograph/pitch meter trace. Calibrations are given in Figure 2, integration time 5 msec.

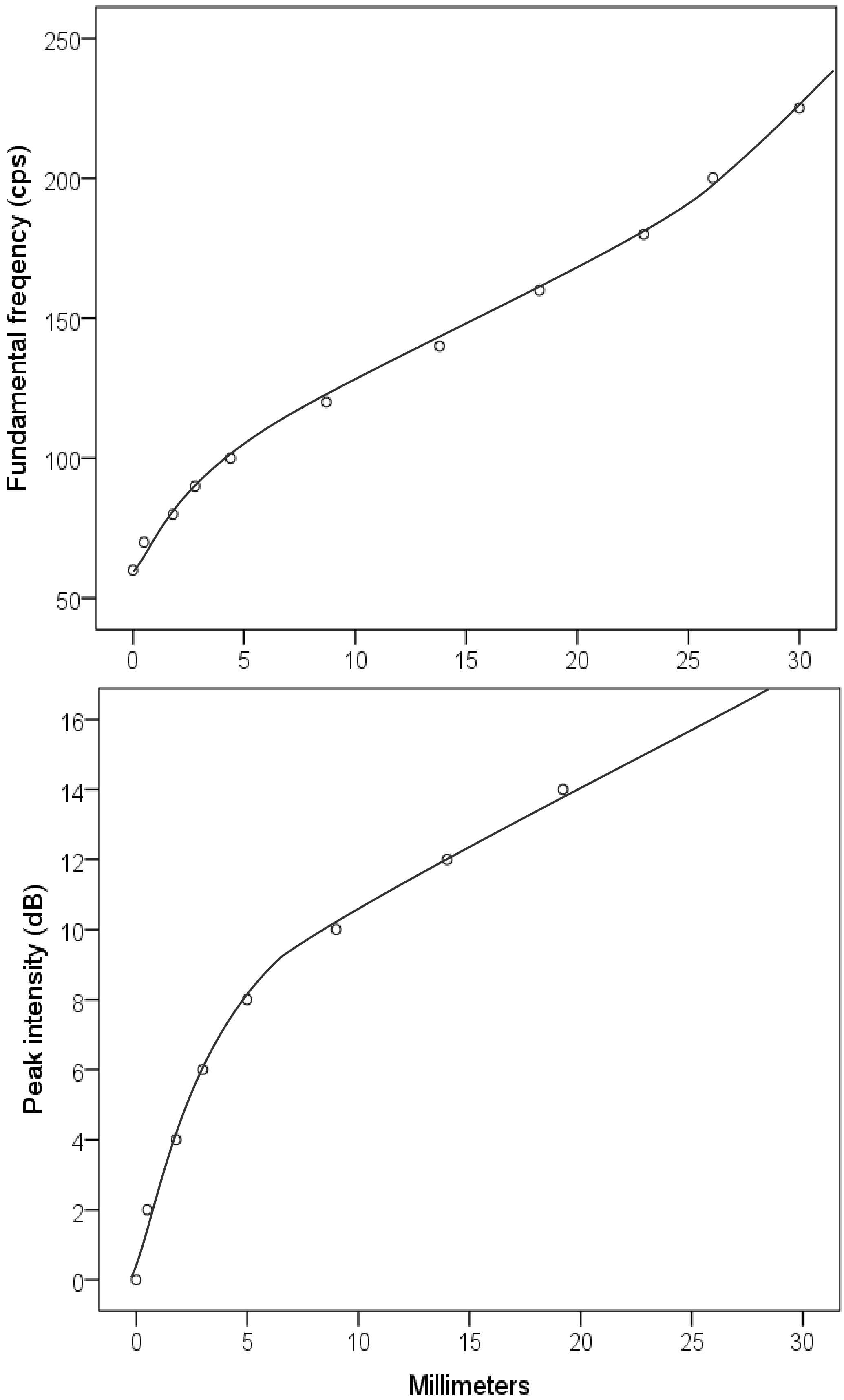
4.4.2 Further analysis

The mingograms were segmented as carefully as possible. The durations of the vowels in the crucial words were measured in csecs. The intensity measurements were based on the peak-intensity values, which were rounded up to the nearest whole decibel.

When F0 was essentially level throughout the vowel, the steady state value was measured. In vowels with falls or rise-falls the highest F0-value was taken as a measure; the lowest value was taken in rises and fall-rises.

The values of these three parameters for the 5 × 35 × 2 vowels are given in Table 2.

To eliminate the influence of the individual speakers we have to concentrate on relative rather than absolute differences between the vowels in each word.



Figures 1 & 2: Calibration of Trans Pitch Meter (top) and Intensity Meter (bottom) used in the acoustic analysis.

Table 2a: Results environment 0_0

Item + speaker	Peak intensity (dB)		F0 (cps.)		Duration (csec.)		Intensity diff (dB)		F0 interval	Duration syll1 (%)
	Syllable		Syllable		Syllable					
	1	2	1	2	1	2	raw	corr.		
unknown										
M1	5	7	117	112	11	9	−2	−3	−1.045	55
M2	9	10	161	175	11	15	−1	−2	−1.087	42
M3	11	11	132	132	9	17	0	−1	1.000	35
M4	10	8	129	156	8	15	2	1	−1.209	35
M5	8	10	155	180	9	13	−2	−3	−1.161	41
eighteen										
M1	14	10	125	130	15	12	4		−1.040	56
M2	15	11	169	200	17	12	4		−1.118	59
M3	15	12	146	158	16	14	3		−1.082	53
M4	12	11	144	168	14	7	1		−1.167	67
M5	19	14	156	192	13	6	5		−1.231	68
mince pie										
M1	10	11	150	155	5	18	−1	1	−1.033	22
M2	8	10	170	194	6	26	−2	0	−1.141	19
M3	12	15	145	163	7	23	−3	−1	−1.124	23
M4	14	10	130	159	5	17	4	6	−1.223	23
M5	9	13	145	168	5	15	−4	−2	−1.159	25
window										
M1	9	7	137	90	9	12	2	5	1.522	43
M2	9	13	174	140	8	12	−4	−1	1.243	40
M3	12	12	150	125	8	15	0	3	1.200	35
M4	13	11	141	107	8	8	2	5	1.318	50
M5	12	13	145	127	6	8	−1	2	1.142	43
footprint										
M1	15	8	184	102	10	7	7		1.804	59
M2	10	7	180	145	10	11	3		1.241	48
M3	13	9	176	123	19	9	4		1.431	68
M4	11	11	176	145	8	7	0		1.214	53
M5	13	12	167	135	7	6	1		1.237	54
absurd										
M1	10	12	137	150	6	15	−2		−1.095	29
M2	10	14	167	185	11	18	−4		−1.108	38
M3	12	15	145	159	11	17	−3		−1.097	39
M4	9	12	133	150	5	14	−3		−1.128	26
M5	12	12	150	176	4	10	−1		−1.173	29
machine										
M1	8	10	127	145	6	7	−2		−1.142	46
M2	6	7	165	127	4	7	−1		1.299	36
M3	8	12	155	175	5	14	−4		−1.129	26
M4	8	10	133	133	4	10	−2		1.000	29
M5	4	7	141	152	6	8	−3		−1.078	43

Table 2b: Results environment 1_0

Item + speaker	Peak intensity (dB)		F0 (cps.)		Duration (csec.)		Intensity diff (dB)		F0 interval	Duration syll1 (%)
	Syllable		Syllable		Syllable					
	1	2	1	2	1	2	raw	corr.		
unknown										
M1	7	7	165	125	8	9	0	−1	1.320	47
M2	4	4	164	180	9	13	0	−1	−1.098	41
M3	9	11	137	137	7	16	−2	−3	1.000	30
M4	12	10	130	150	8	12	2	−1	−1.154	40
M5	10	13	166	160	9	14	−3	−4	1.038	39
eighteen										
M1	12	8	135	134	12	11	4		1.008	52
M2	13	9	176	194	17	12	4		−1.102	59
M3	12	11	168	162	17	18	1		1.037	49
M4	11	11	163	180	13	8	0		−1.104	62
M5	10	11	167	193	11	8	−1		−1.142	58
mince pie										
M1	6	9	140	130	6	16	−3	−1	1.077	27
M2	8	8	170	180	7	25	0	2	−1.065	22
M3	11	15	163	195	6	27	−4	−2	−1.296	18
M4	13	11	145	173	7	22	2	4	−1.193	24
M5	11	12	159	163	5	15	−1	1	−1.025	25
window										
M1	6	3	139	85	6	14	3	6	1.635	30
M2	4	8	167	90	8	18	−4	−1	1.856	31
M3	10	9	145	121	6	8	1	4	1.199	43
M4	13	9	139	102	6	8	4	7	1.363	43
M5	13	15	150	127	7	7	−2	1	1.181	50
footprint										
M1	11	6	159	90	8	8	5		1.767	50
M2	8	7	224	120	11	8	1		1.867	58
M3	10	10	180	115	10	7	0		1.565	59
M4	13	12	176	156	10	7	1		1.128	59
M5	15	12	175	145	8	5	3		1.207	62
absurd										
M1	9	10	156	119	10	13	−1		1.311	43
M2	6	12	187	212	6	19	−6		−1.134	24
M3	4	15	150	164	8	17	−9		−1.093	32
M4	8	9	147	150	8	13	−1		−1.020	38
M5	9	12	167	129	5	11	−3		1.295	31
machine										
M1	7	1	128	128	5	5	−3	−6	1.000	50
M2	9	8	140	123	8	8	1	−2	1.138	50
M3	11	13	145	163	6	14	−2	−5	−1.124	30
M4	12	10	133	128	7	9	2	−1	1.039	44
M5	10	12	143	147	5	8	−2	−5	−1.028	38

Table 2c: Results environment 0_1

Item + speaker	Peak intensity (dB)		F0 (cps.)		Duration (csec.)		Intensity diff. (dB)		F0 interval	Duration syll1 (%)
	Syllable		Syllable		Syllable					
	1	2	1	2	1	2	raw	corr.		
unknown										
M1	12	9	138	138	10	12	3	2	1.000	45
M2	8	9	150	150	12	13	−1	−2	1.000	48
M3	15	12	140	140	10	15	3	2	1.000	40
M4	13	10	129	129	8	12	3	2	1.000	40
M5	15	13	157	157	9	10	2	1	1.000	47
eighteen										
M1	12	8	137	150	14	9	4		−1.095	61
M2	15	10	205	194	15	12	5		1.057	56
M3	15	11	172	163	13	12	4		1.055	52
M4	14	10	167	137	12	11	4		1.219	52
M5	15	10	163	165	19	9	5		−1.012	68
mince pie										
M1	9	11	127	103	5	20	−2	0	1.233	20
M2	6	9	160	163	7	20	−3	−1	−1.019	26
M3	10	13	154	158	8	25	−3	−1	−1.026	24
M4	12	8	140	131	6	23	4	6	1.069	21
M5	12	13	165	159	6	14	−1	1	1.038	30
window										
M1	11	10	137	90	18	13	1	4	1.522	58
M2	12	11	170	130	9	16	1	4	1.307	36
M3	11	9	147	118	10	16	2	5	1.246	38
M4	14	12	129	121	5	14	2	5	1.066	26
M5	11	14	152	129	7	8	−3	0	1.178	47
footprint										
M1	11	7	174	103	10	6	4		1.689	63
M2	13	7	225	121	11	10	6		1.850	52
M3	12	11	180	145	9	6	1		1.241	60
M4	15	12	186	145	9	5	3		1.283	64
M5	15	13	197	137	8	6	2		1.438	57
absurd										
M1	11	12	102	132	9	15	−1		−1.294	38
M2	7	13	157	205	7	21	−6		−1.306	25
M3	8	11	130	156	5	15	−3		−1.200	25
M4	6	11	123	161	4	18	−5		−1.309	18
M5	9	13	167	215	4	13	−4		−1.287	24
machine										
M1	7	10	137	150	6	10	−3	−6	−1.095	38
M2	4	3	175	193	6	11	1	−2	−1.103	35
M3	9	11	150	167	6	13	−2	−5	−1.113	32
M4	15	14	129	141	6	18	1	−2	−1.093	25
M5	11	12	141	167	4	8	−1	−4	−1.184	33

Table 2d: Results environment 1_1

Item + speaker	Peak intensity (dB)		F0 (cps.)		Duration (csec.)		Intensity diff (dB)		F0 interval	Duration syll1 (%)
	Syllable		Syllable		Syllable					
	1	2	1	2	1	2	raw	corr.		
unknown										
M1	8	8	160	160	8	12	0	−1	1.000	40
M2	9	10	170	170	9	14	−1	−2	1.000	39
M3	9	10	163	143	8	15	−1	−2	1.140	35
M4	12	12	135	135	10	13	0	−1	1.000	43
M5	12	13	174	161	8	14	−1	−2	1.081	36
eighteen										
M1	15	10	146	155	18	8	5		−1.054	69
M2	13	10	213	194	17	11	3		1.098	61
M3	13	11	165	145	14	12	2		1.138	54
M4	14	9	150	149	14	10	5		1.007	58
M5	19	15	187	200	11	11	4		−1.070	50
mince pie										
M1	10	13	164	173	5	17	−3	−1	−1.055	23
M2	6	9	182	163	6	19	−3	−1	1.117	24
M3	11	14	144	161	7	22	−3	−1	−1.117	24
M4	12	11	142	160	6	16	1	3	−1.160	27
M5	10	11	167	120	4	13	−1	1	1.392	24
window										
M1	7	5	145	90	5	15	2	5	1.611	25
M2	8	11	175	115	6	14	−3	0	1.522	30
M3	10	13	145	118	9	19	−3	0	1.229	32
M4	13	9	137	100	5	11	4	7	1.370	31
M5	14	15	162	121	5	8	−1	2	1.339	38
footprint										
M1	3	1	150	90	8	7	2		1.667	53
M2	13	7	250	125	12	12	6		2.000	50
M3	9	9	167	137	10	7	0		1.219	59
M4	11	8	156	137	7	6	3		1.139	54
M5	15	13	167	155	7	5	2		1.077	58
absurd										
M1	10	12	167	129	6	17	−2		1.295	26
M2	5	13	163	193	4	10	−8		−1.184	29
M3	5	15	145	156	8	15	−10		−1.076	35
M4	11	13	143	154	6	12	−2		−1.077	33
M5	14	15	193	176	4	11	−1		−1.097	27
machine										
M1	9	10	155	167	6	8	−1	−4	−1.077	43
M2	9	7	158	125	7	12	2	−1	1.264	37
M3	11	13	145	176	5	17	−2	−5	−1.152	23
M4	10	6	129	156	6	12	4	1	−1.209	33
M5	5	6	141	138	4	9	−1	−4	1.022	31

Table 2e: Results environment #_#

Item + speaker	Peak intensity (dB)		F0 (cps.)		Duration (csec.)		Intensity diff (dB)		F0 interval	Duration syll1 (%)
	Syllable		Syllable		Syllable					
	1	2	1	2	1	2	raw	corr.		
unknown										
M1	9	7	120	120	8	12	2	1	1.000	40
M2	9	10	160	175	11	15	−1	−2	−1.094	42
M3	11	14	135	135	11	17	−3	−4	1.000	39
M4	12	13	117	125	8	15	−1	−2	−1.068	35
M5	13	14	155	155	10	17	−1	−2	1.000	37
eighteen										
M1	11	6	131	137	15	9	5		−1.046	63
M2	10	9	153	193	17	18	1		−1.261	49
M3	10	8	125	148	15	14	2		−1.184	52
M4	15	12	125	140	14	19	3		−1.120	42
M5	15	13	159	159	14	14	2		1.000	50
mince pie										
M1	9	6	127	130	5	11	3	5	−1.024	31
M2	10	6	148	160	8	16	4	6	−1.081	33
M3	13	11	132	157	4	30	2	4	−1.189	12
M4	15	10	122	150	8	32	5	7	−1.230	20
M5	15	15	159	156	6	26	0	2	1.019	19
window										
M1	10	6	130	85	8	14	4	7	1.529	36
M2	12	11	176	134	8	15	1	4	1.313	35
M3	14	12	145	105	10	22	2	5	1.381	31
M4	14	9	120	90	10	19	5	8	1.333	34
M5	15	15	168	106	6	10	0	3	1.585	38
footprint										
M1	14	4	145	97	8	7	10		1.495	53
M2	14	4	215	117	10	8	10		1.838	56
M3	11	8	147	125	12	10	3		1.176	55
M4	15	9	145	102	7	8	6		1.422	47
M5	16	14	192	101	7	7	2		1.901	50
absurd										
M1	4	9	102	107	6	27	−5		−1.049	18
M2	4	14	127	193	11	24	−10		−1.520	31
M3	2	14	117	145	11	21	−12		−1.239	34
M4	10	15	112	138	11	24	−5		−1.232	31
M5	13	17	143	192	8	26	−4		−1.343	24
machine										
M1	4	7	110	160	6	11	−3	−6	−1.455	35
M2	9	12	147	213	7	13	−3	−6	−1.449	35
M3	11	13	131	154	6	19	−2	−5	−1.176	24
M4	11	13	117	143	5	14	−2	−5	−1.222	26
M5	11	13	150	177	8	15	−4	−7	−1.180	35

4.4.2.1 Relative durational differences

The duration values are given in percentages which represent the duration of the first vowel as proportional to the total duration of the two vowels in a particular word when added together. In this way it is possible to eliminate the influence of individual differences in tempo.

4.4.2.2 Relative F0-differences

The relative F0-differenees are expressed in what I have called an interval index. This index is calculated by dividing the higher cps.-value by the lower one, which yields an index between 1.000 and − as no interval greater than one octave was found in the corpus − 2.000. For the sake of comparison I have included Table III containing interval indices for 1, 2, 3, ... 12 semitones. The ratios on which these indices are based are taken from Helmholtz (1954: 17). The F0 interval can be computed in semitones from the index by taking the logarithm to the base 2 and multiplying the result by 12: $12 \times {}^{2}\log(\text{index})$.

Table 3: Interval indices for 0-12 semitone intervals

Number of semitones	Ratio	Interval index
0	1 : 1	1.000
1	9 :10	1.111
2	8 : 9	1.125
3	5 : 6	1.200
4	4 : 5	1.250
5	3 : 4	1.333
6	5 : 7	1.400
7	2 : 3	1.500
8	5 : 8	1.600
9	3 : 5	1.666
10	4 : 7	1.750
11	5 : 9	1.800
12	1 : 2	2.000

4.4.2.3 Intensity differences

Uncorrected intensity differences.

The relative intensity differences are found by simply subtracting the one dB-value from the other.

All data concerning durational proportions, F0 and intensity differences are tabulated in Table II. In these tables a positive value means that the first syllable in a word has the higher value of the two; a negative value means that the second syllable is the stronger.

Intensity differences corrected for inherent sonority.

It has been suggested earlier in this paper, as well as in the literature, that the contribution of inherent sonority to the total intensity of a particular vowel may very well be an irrelevant factor in the perception of stress. Conversely, I would argue that the specification of an acoustic basis of stress patterns is obscured by inherent sonority. It happened e.g. that the second vowel in a word like *window* had a greater intensity than the first, although the stress was on the first syllable. By correcting the vowels for inherent sonority, the balance might be restored to a proper falling stress. The correction procedure actually used in this experiment was the following: all the vowels in the crucial words were given an extra intensity as if they all had the phonetic quality of the vowel a. The correction factors that were used are given in Table 4. They are, in fact, the factors suggested by Lehiste and Peterson (1958: table x, row iii) rounded up to integral decibels. Thus, in those cases where the correction factor was less than .5 dB, no correction was applied at all.

Table 4: Intensity corrections for inherent sonority.

Word	Suggested correction factor
unknown	1 dB extra on second syllable
eighteen	no correction
mince pie	2 dB extra on first syllable
window	3 dB extra on first syllable
footprint	no correction
absurd	no correction
machine	3 dB extra on second syllable

4.4.3 Averaging

The duration proportions, F0-indices, corrected and uncorrected intensity differences were averaged over the five individual speakers. The average values are tabulated separately in Table 5.

The individual as well as the averaged values are presented graphically in Figure 3 (duration proportions), Figure 4 (F0-intervals), Figure 5 (uncorrected intensity difference) and Figure 6 (corrected intensity differences). In these figures the data are grouped by phono-syntactic environment, and differentiated for each of the seven words.

4.5 Conclusions and discussion

4.5.1 Identification of stress patterns on an acoustic basis

4.5.1.1 Duration proportions

First of all, it must be obvious that the duration proportions cannot be used to compare among words. Each vowel in English has its own typical length, and no attempts have been made here to correct for inherent length. Although such correction factors have been tentatively proposed by Peterson and Lehiste (1960, table i) I did not consider it worthwhile following it up, as no systematic differences can be detected among e.g. the (1-1), (2-1), and (1-2) stress patterns of double stressed words, which are different on account of the rhythmic principle.

4.5.1.2 Intensity differences

It seems to me that uncorrected intensity differences can effectively distinguish the (3-1) words from the other types in the #_# context, and marginally in other contexts. When corrected for inherent sonority, intensity differences become discriminatory for all contexts, at least with respect to the (3-1) stress pattern. (cf. Figures 5 and 6).

4.5.1.3 Fundamental frequency indices

The (1-2) or falling stress pattern can easily be isolated in all phono-syntactic environments on the basis of a +.5 interval index.

In the 0_0 and the 1_0 contexts the F0 interval is not discriminatory between the words with level or rising stress. In the three remaining environments this distinction can be made, where the index is about −.1 for the double stressed words, and about −.2 for the (3-1) type.

4.5.1.4 Combination of factors

It seems to me that we can effectively recognize three patterns on the basis of a combination of cues: the (1-2) pattern can always be identified by its considerably positive F0-index; a further distinction can be drawn between the (3-1) patterns and the double stressed words on the basis of intensity differences, especially when these are corrected for inherent sonority. In three out of five contexts, however, we can dispense with this cue, as the F0-index is powerful enough by itself.

Table 5: F0-intervals (index, semitones), corrected and uncorrected intensity differences (dB) and duration proportions, averaged over the five subjects.

Word + context	F0_index	F0 interval (st)	Int. dif_raw	Int. dif_cor	Dur%_1
unknown					
0_0	−1.082	−1.31	−.6	−1.6	41.6
1_0	1.122	0.27	−.6	−2.0	39.4
0_1	1.000	0.00	2.0	1.0	44.0
1_1	1.044	0.72	−.6	−1.6	38.6
#_#	−1.032	−0.54	−.8	−1.8	38.6
Eighteen					
0_0	−1.128	−2.25	3.4		60.6
1_0	−1.061	−1.03	1.6		56.0
0_1	1.049	0.71	4.4		57.8
1_1	1.024	0.35	3.8		58.4
#_#	−1.122	−1.94	2.6		51.2
mince pie					
0_0	−1.136	−2.18	−1.2	.8	22.4
1_0	−1.084	−1.26	−1.2	.8	23.2
0_1	1.059	0.93	−1.0	1.0	24.2
1_1	1.089	0.54	−1.8	.2	24.4
#_#	−1.101	−1.60	2.8	4.8	23.0
window					
0_0	1.285	4.25	−.2	2.8	42.2
1_0	1.447	6.12	.4	3.4	39.4
0_1	1.264	3.93	.6	3.6	41.0
1_1	1.414	5.92	−.2	2.8	31.2
#_#	1.428	6.12	2.4	5.4	34.8
Footprint					
0_0	1.385	5.44	3.0		56.4
1_0	1.507	6.75	2.0		57.6
0_1	1.502	6.83	3.2		59.2
1_1	1.620	5.56	2.6		54.8
#_#	1.566	7.50	6.2		52.2
Absurd					
0_0	−1.120	−1.96	−2.6		32.2
1_0	1.072	1.02	−4.0		33.6
0_1	−1.279	−4.26	−3.8		26.0
1_1	−1.028	0.12	−4.6		30.0
#_#	−1.277	−1.96	−7.2		27.6
Machine					
0_0	−1.101	−0.23	−2.4	5.4	36.0
1_0	1.005	0.08	−.8	−3.8	42.4
0_1	−1.118	−1.92	−.8	−3.8	32.6
1_1	−1.039	−0.70	.4	−2.6	33.4
#_#	−1.296	−4.41	−2.8	−5.8	31.0

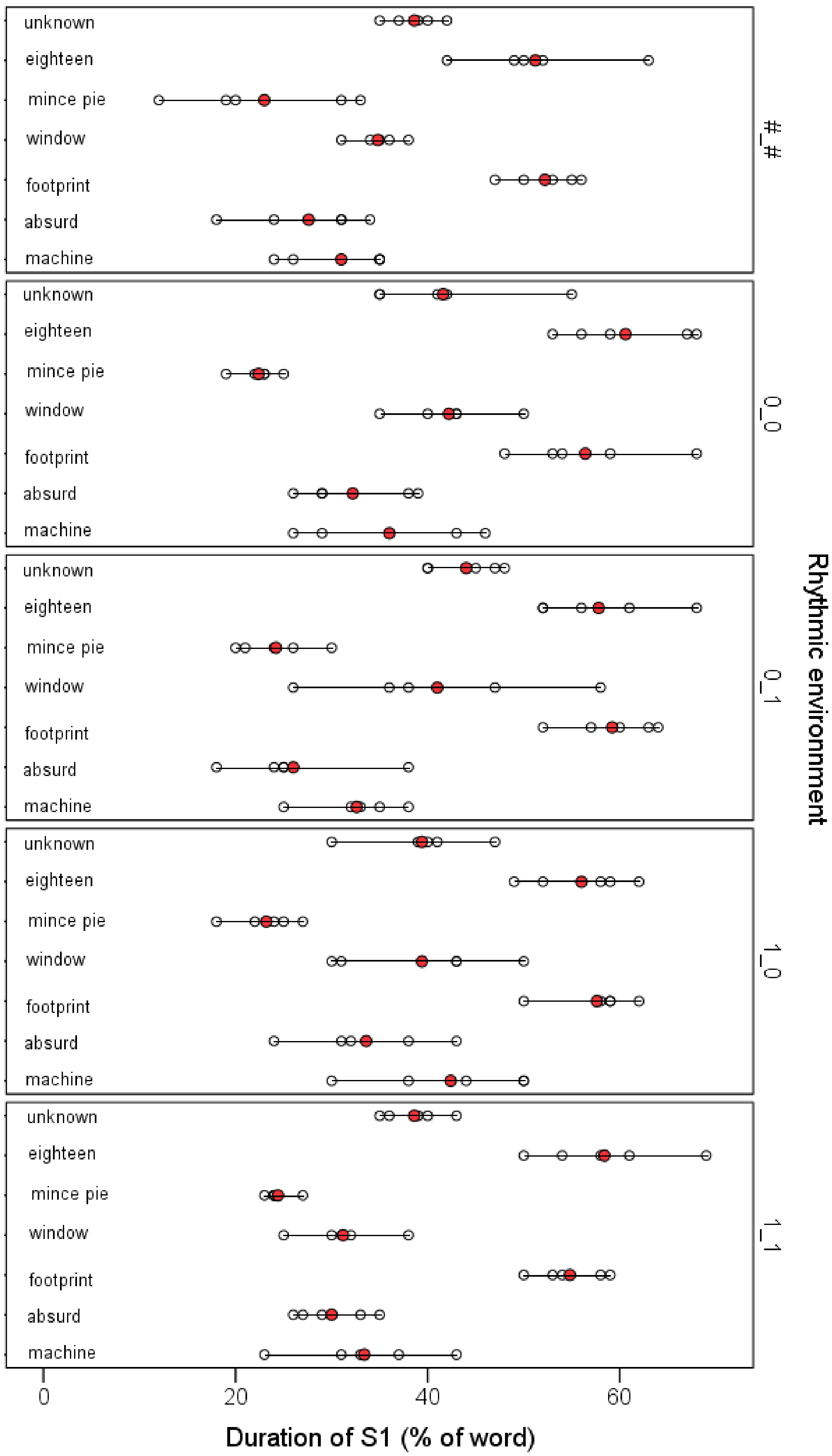


Figure 3: Relative duration of first syllable (% of word length) broken down by rhythmic environment and target word.

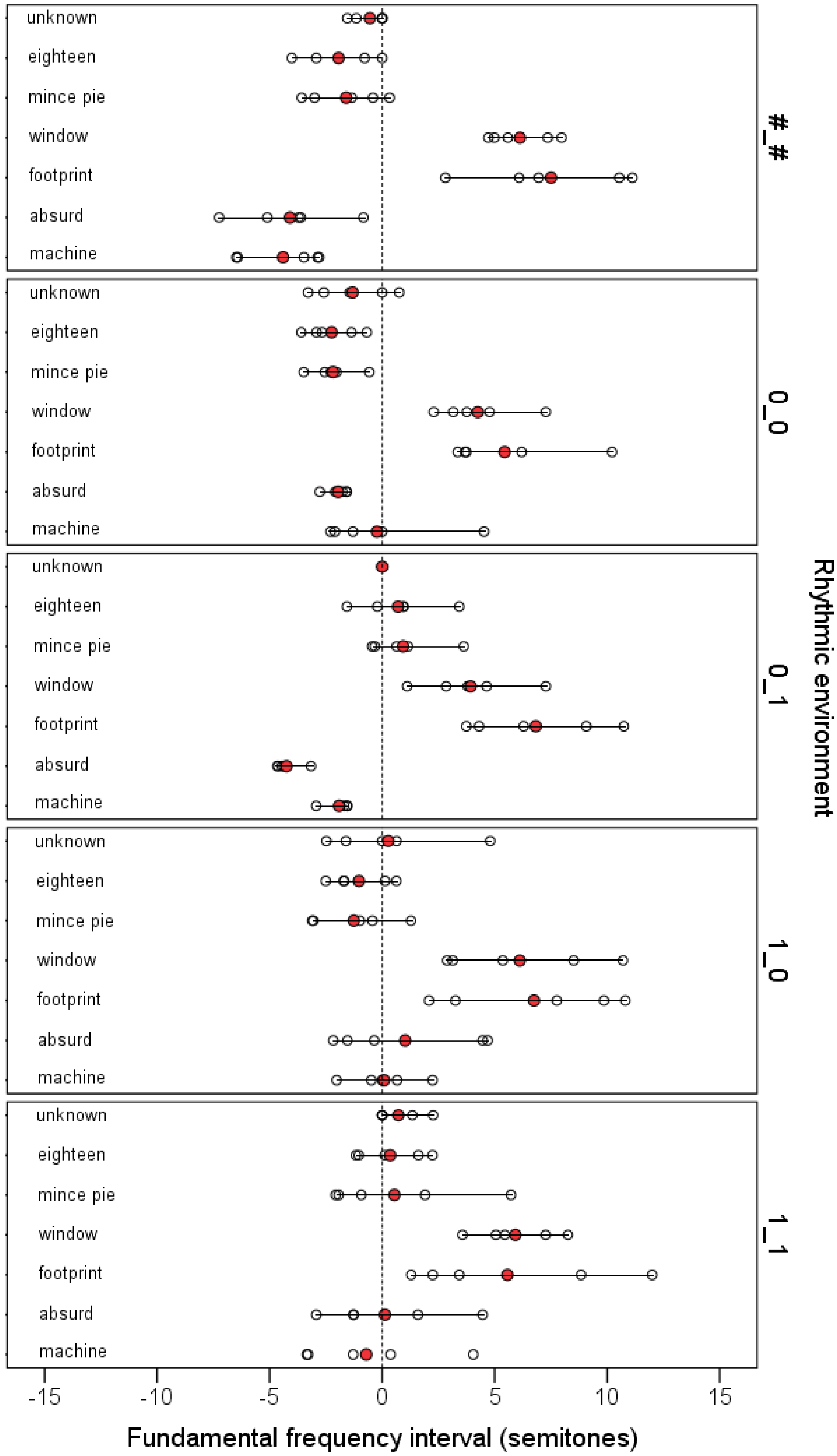


Figure 4: Fundamental frequency difference between first and second syllable (semitones) broken down by rhythmic environment and target word.

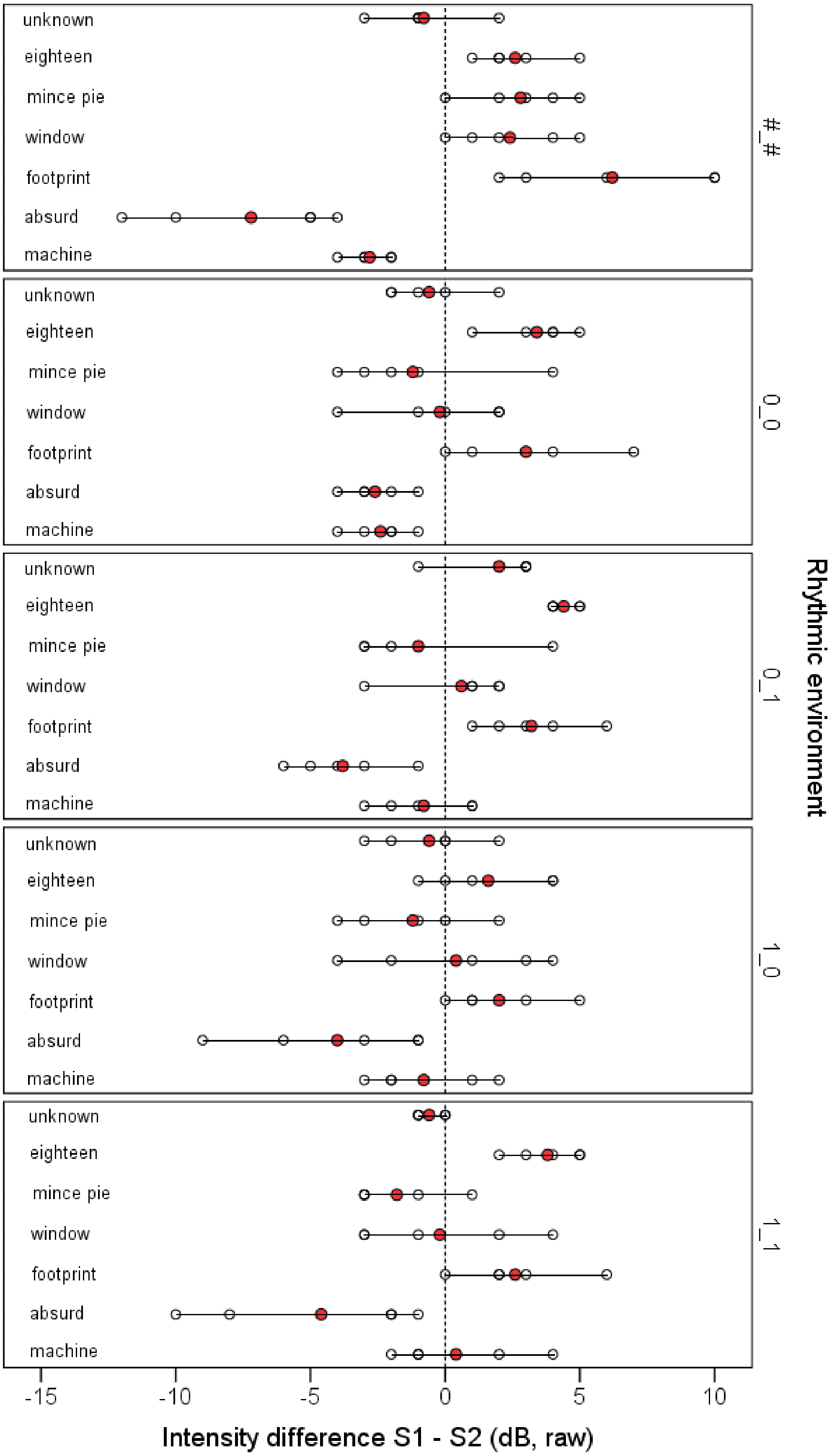


Figure 5: Intensity difference between first and second syllable (decibels, not corrected for inherent vowel intensity) broken down by rhythmic environment and target

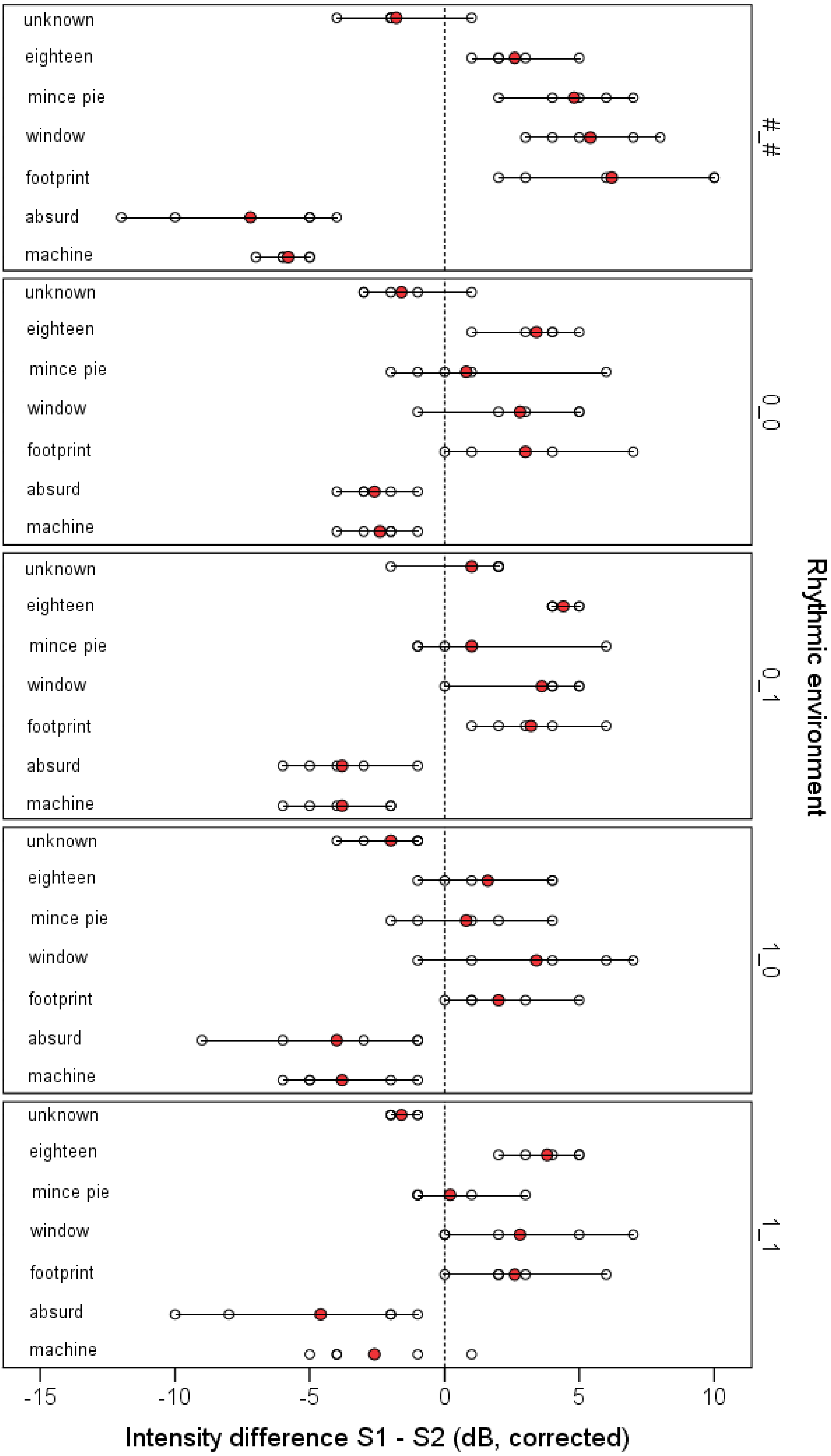


Figure 6: Intensity difference between first and second syllable (decibels, corrected for inherent vowel intensity) broken down by rhythmic environment and target word.

4.5.2 Acoustic evidence for double stress

We have identified the (3-1) and (1-2) patterns, and a group of words somewhere in between these two. So far I have avoided the question if a further distinction in this middle class is possible. As anticipated in chapter II, a positive answer to this question would be a strong indication that the traditional handbooks were essentially correct in postulating a (1-1) pattern along side the (2-1) pattern. This question can be settled by considering the effects of the rhythmic principle. As explained in my introductory chapter, a (1-1) pattern is expected in 0_0 contexts, (1-2) in 0_1, and (2-1) in 1_0.

There is evidence that the shift from (1-1) to (1-2) is real, as the F0-index shifts from slightly negative to slightly positive, viz. from −.1 to +.1. Since this effect of the rhythmic principle has never been questioned, this result is in not very surprising.

There is, however, no difference in terms of F0 indices between the 0_0 and 1_0 realizations. In both oases the indices are approximately −.1.

But, as I have said en an earlier occasion, stress patterns had rather be considered in relation to each other than from an absolute view point. The neighbouring falling stress pattern (1-2) has a +.3 F0-index in the 0_0 context, but about +.5 when realized in a 1_0 environment. Thus the distance between the typical F0-index for (1-1) and (1-2) patterns is about .4 under 0_0, and about .6 under 1_0 circumstances. An interval index of +.5 is also the typical value for all other contexts. Therefore I suggest the following: the fact that under a 1_0 condition the distance to the neighbouring stress pattern is increased remains perceptually unnoticed, and the situation is interpreted as if the stress on the first syllable of a double stressed word is lowered instead.

A further implication of this view is that all realizations of double-stressed words in 1_1 and #_# contexts are to be interpreted as (2-1) patterns, as the difference between this pattern and the (1-2) pattern in terms of F0 interval index is also about .6. This is supported by two facts: firstly, the 1_1 and 1_0 contexts exert the same influence on double stressed words (Kingdon 1958b: 165; van Heuven 1973: 29); secondly, it has been claimed in the literature (Vanvik 1962: 66) that no (1-1) realization of double-stressed words is possible in citation forms.

Summing up I would say that there are reasons to believe that double stress exists, but only in contexts where no stressed syllable precedes or follows the double stressed word. Double stress cannot be acoustically characterized in any absolute sense but it can be separated from other stress patterns when the patterns are considered from a relative angle. Finally, it must be apparent that stress patterns can be correlated with acoustic parameters.

4.5.3 Implications for synthetic stimuli

It stands to reason that the amount of variation in synthetic stimulus material should not be different from what happens in natural speech. Thus I decided not to vary the duration of the syllables in the crucial words.

A second implication is that the F0 interval indices must range between +.5 and (partly for the sake of symmetry) −.5, with some typical intermediate values at +.3, +.1, −.1, −.3.

It has appeared that intensity differences have considerable discriminatory power, so these will have to be incorporated in the synthetic stimuli as well. The bounds of variation are typically +5 dB and −5 dB.

Chapter five

Central experiments: Synthesis

5.0 Introduction

Throughout this investigation I have made the assumption that stress, and especially stress levels and patterns, are ultimately perceptual phenomena. This means that only perceptual evidence can supply an answer to our problem. This assumption is reasonable, and is basic to much recent work on stress.

The question whether or not double stress exists in English therefore had to be answered by a perception test. To this effect I devised a test which had as its stimulus material a number of sentences, each of which contained one crucial word. Crucial words were either of the double stress, rising stress, or falling stress type. For each several prosodically different versions were synthesized by systematically altering the acoustic make-up of the crucial words. In this way sets of sentences were created which had a range of stress patterns on the crucial words varying from extremely rising stress, through a level distribution of stress over the two syllables, to extremely falling. Before this material could be synthesized, a number of questions had to be solved:

(1) How, i.e. by altering what acoustic parameters can we effectively create the perceptual impression of a variety of stress patterns?
(2) What is the optimal range of these parameters needed to create these effects?
(3) How many intermediate steps do we need to cover the range between the extremes?
(4) How do we know that the steps are close enough to each other to sample the range adequately?
(5) How do we know that the steps are big enough to be auditorily distinct?

The answers to the first two questions have already been given in § 4.5.3; it should be noted in this context that I have limited myself to two parameters for purely practical reasons.

The number of steps in the stress dimension was more or less axiomatically set at seven, where three rising and three falling patterns were placed symmetrically round the middle pattern at an exact equilibrium of the parameter values. The non-level patterns were synthesized so as to approximate the typical parameter values of the stress patterns identified in the previous chapter.

The answer to the last question could not be given without the aid of some more experiments, viz. pretests 1 and 2.

5.1 Pretest 1

5.1.0 Introduction

Seven prosodically different versions of a sentence were prepared which were identical to the ones to be used in the main test in virtually every respect. The objective of the pretests was to see if the subjects could tell these seven versions apart. Their discriminatory ability was tested by having them perform two different tasks. In this section I will report on the first of these tasks.

Since the synthesized material was in many ways similar to that in the main test, it stands to reason that I will outline the synthesis procedure only once, here, and refer to this section on all further occasions.

5.1.1 Stimuli

5.1.1.1 Choice of basic material

As a carrier the structure *Are they .... in 'your country?* was chosen. I have opted for an interrogative form as this is often advocated in the literature (Lehiste and Peterson 1958; Lieberman 1967; Lehto 1969). Also, in an experiment carried out by myself I obtained better results with question forms than with assertions (van Heuven 1972).

The stress for emphasis on the word *'your* was included to remove sentence stress from whatever was to be inserted on the dots. Vanvik (1962: 67) has led me to suspect that double stresses are not very likely to occur under sentence stress. My own findings (§ 4.5.2) tend to corroborate this.

Though a variety of words and combinations of words were to be inserted on the dots in the carrier for the main experiment, I decided that the nonsense word *sisis* would be sufficient, in fact more suitable, for the purpose of the pretests. Since this word is meaningless, listeners do not expect any particular stress pattern on it, so that their stress perception will be entirely motivated by the acoustic make-up of the signal. Secondly, it serves to eliminate the question of inherent sonority from this set up, as the two constituent syllables of this word are identical (cf. Morton and Jassem 1965: 163).

5.1.1.2 Synthesis

The utterance *Are they sisis in 'your country?*, spoken by a male speaker of R.P. English, was recorded on tape and its fundamental frequency curve was drawn according to narrow-band spectrogram tracings of the third harmonic.

The same sentence was then synthesized on PAT, an eight-parameter acoustic analogue of the human vocal tract (Lawrence 1953) at Edinburgh University. The parameters were controlled by a punched paper tape on which the values for each of the parameters were stored digitally per 10 msecs. This tape was punched by a computer on the basis of an R.P. English synthesis-by-rule program which was essentially the same as the one described by Holmes, Mattingly and Shearme (1964). By rule of thumb this program gives all vowels the same fundamental amplitude (A0), but no provisions are incorporated for fundamental frequency. Therefore the F0 information of the narrow-band spectrogram tracings was separately given to the computer. The F0, however, was kept constant over the whole duration of the word *sisis*, and as such linked the F0 levels of the neighbouring left and right-hand sounds by the best fitting straight line.

The result was my basic, prosodically level, version. Six deviating versions were then synthesized by varying the A0 and F0 values of the two vowels in the crucial word *sisis*, leaving the rest of the sentence intact.

The A0 excursions from the reference levels in the basic sentence were in steps of 1.75 dB, which is the smallest step that can be handled by the computer program. The F0 steps were very small as well, though not always as small as possible.

When the parameter value would be a step up from the reference level in the first syllable the value for the second syllable would be decreased simultaneously, and *vice versa*, so that the difference between the two syllables was twice as big as the step. When A0 was increased, F0 was increased, such that smallest A0 excursions were paired with the smallest F0 excursions, and that progressively larger A0 and F 0 excursions went together. The size of the excursions was chosen so as to optimally approximate the values for the various stress patterns that were stipulated earlier in this report in § 4.5.3.

Complete information concerning the 7 versions is given in Table 6 for A0 and Table 7 for F0, and represented graphically in Figure 7. To give the reader an indication of the overall acoustics of the stimulus material I have included various spectrograms in Appendix II; Appendix III, finally, is the printed-out version of the control punched paper tape and a conversion table to translate the levels 1 to 31 to acoustic measures (see also: Holmes, Mattingly and Shearme 1964: appendix).

5.1.1.3 Arrangement of stimuli

For convenience of reference I have given the following names to the seven synthesized stress patterns:

rising stress, extreme:	(−3)
rising stress, intermediate:	(−2)
rising stress, slight:	(−1)
level stress:	(0)
falling stress, slight:	(+1)
falling stress, intermediate:	(+2)
falling stress, extreme:	(+3)

This symbolization will be used in other sections further on in this report as well as in the relevant figures, tables, and appendices. The synthesized sentences were recorded in pairs, such that the two members of each pair were always contingent in terms of stress differences, e.g. (−3, −2), (−2, −1), (−1, 0), (+2, +3), etc. In some instances the first member of a pair was the more extreme stress pattern; in other cases the situation was reversed. The pairs (−1, 0) and (+1, 0), where the acoustic differences between the pairs are slighter than in any other combination, were recorded both ways.

Each of the resulting ten pairs was recorded twice, and each item, consisting of two pairs, was preceded by its item number. The complete items were interspaced at 10-sec. intervals.

Table 6: Fundamental amplitude (A0) variations in the target words.

Stress pattern	Program level		Intensity (dB)		Difference (dB)
	Syll. 1	Syll. 2	Syll. 1	Syll. 2	
(0)	29	29	50.75	50.75	0.00
(−1)	29	30	50.75	52.50	−1.75
(−2)	29	31	50.75	54.25	−3.50
(−3)	28	31	49.00	54.25	−5.25
(+1)	30	29	52.50	50.75	1.75
(+2)	31	29	54.25	50.75	3.50
(+3)	31	28	54.25	49.00	5.25

Table 7: Fundamental frequency (F0) variations in the target words.

Stress patterns	Program level		F0 (c.p.s.)		F0 interval index	
	Syll. 1	Syll. 2	Syll. 1	Syll. 2	Obtained	Aimed at
(0)	16	16	120	120	1.000	1.000
(−1)	15	17	116	126	–1.086	–1.100
(−2)	13	18	103	132	–1.282	–1.300
(−3)	12	20	97	146	–1.505	–1.500
(+1)	17	15	126	116	1.086	1.100
(+2)	18	13	132	103	1.282	1.300
(+3)	20	12	146	97	1.505	1.500

Figure 7-a: PAT **s**ynthesis parameters for carrier sentence and base version of nonsense word. Parameter values are between 0 and 31. For conversion of parameter levels to physical units see Appendix III.

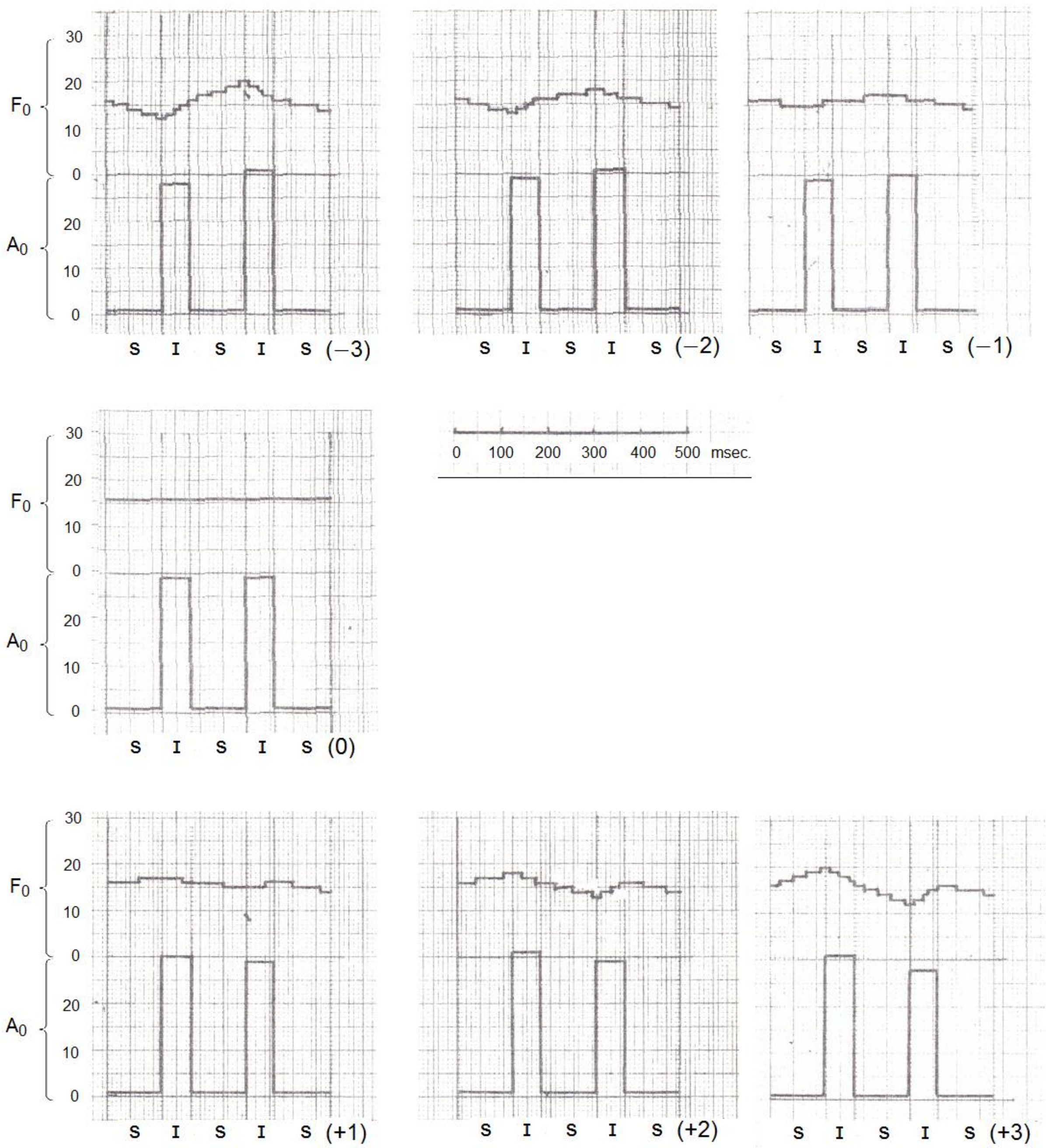
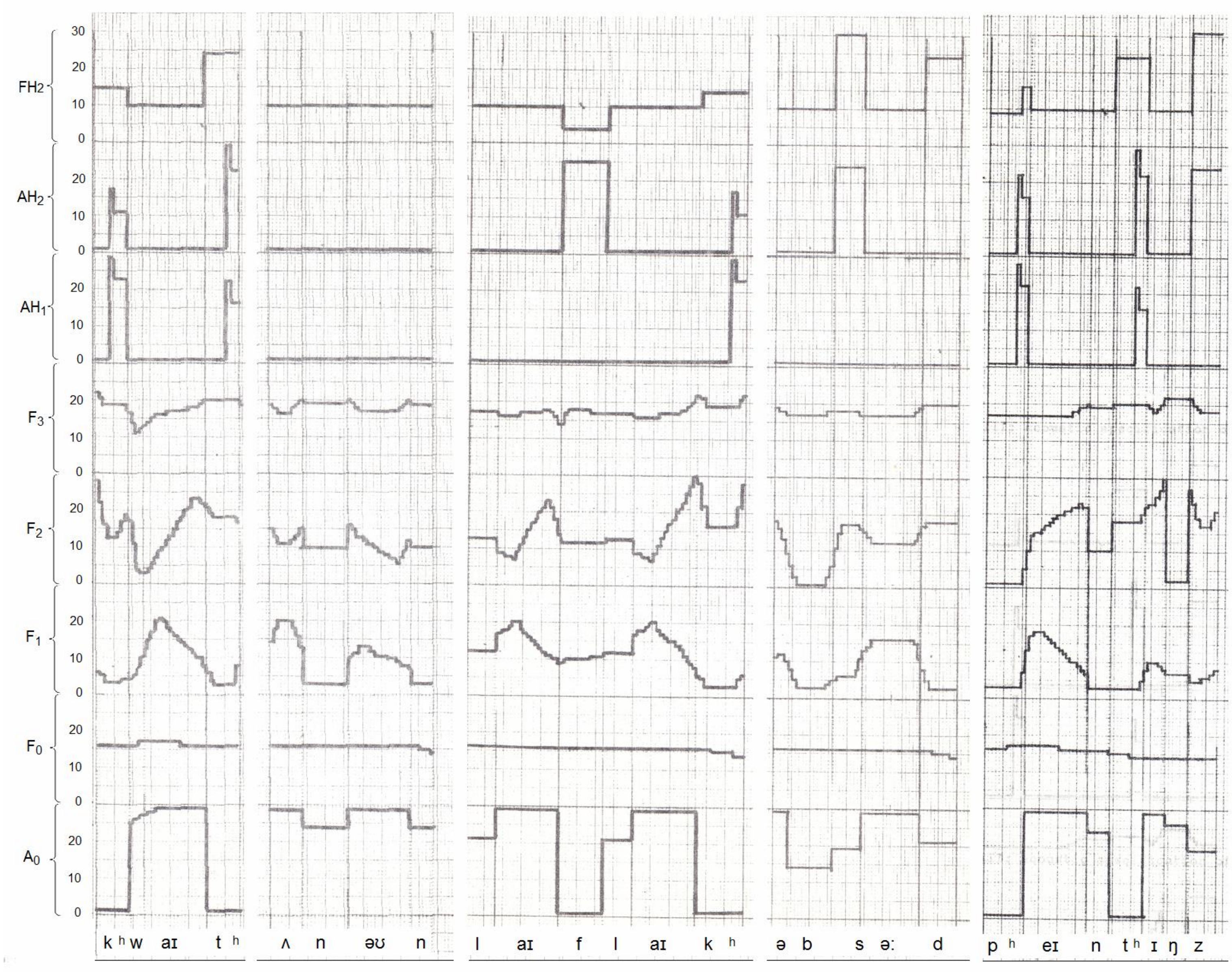


Figure 7-b: PAT **s**ynthesis parameters (F0 and A0) for seven stress patterns on nonsense word inserted in carrier sentence.

Figure 8: (next page) PAT **s**ynthesis parameters for target words to be inserted in carrier sentence. Parameter values are between 0 and 31. For conversion of parameter levels to physical units see Appendix III.


FH2
AH2
AH1
F3
F2
F1
F0
A0
30
20
10
0
kʰ w aɪ tʰ
ʌ n əʊ n
l aɪ f l aɪ kʰ
ə b s əː d
pʰ eɪ n tʰ ɪ ŋ z

5.1.2 Subjects

Twenty-seven subjects took part in the experiment. They were students and staff, male and female, at Edinburgh University. As they were (applied) linguists, phoneticians and speech therapists, none of them can be called linguistically naive.

All subjects were native speakers of some variety of British English. Most of them were professed R.P. speakers, but in a number of cases they were regional dialect speakers, as there were simply not enough R.P. speakers available at the time.

5.1.3 Procedure

The subjects were issued with written instructions and answer sheets, specimens of which are included in Appendix IV. They were required to listen to the tape carefully, and then decide whether the first or the second member of a pair was more extremely characterized for stress, and to indicate their choice on the answer sheets. They were to gamble in case of doubt.

The test was conducted in an ordinary class-room situation, though precautions had been taken that the subjects were placed at roughly equal distances from the loudspeakers of the tape recorder.

The tape was played twice, once to let the listeners get accustomed to synthetic speech, and to give them a general idea of what their task was, and the second time as the test proper.

5.1.4 Results, analysis, and conclusions

Table 8 contains specifications of each of the 8 stimuli, their order in the presentation, and the subject's reactions to them. Had there been no audible differences between the stress patterns of the crucial words in the items, about half of the subjects would have guessed that the first of the two patterns was the more extreme, the other half would have voted for the second pattern. Only if the number of subjects in favour of one particular choice is sufficiently greater than 50%, we can conclude that one of the two patterns must have audibly more extremely stressed.

Table 8: Results pretest 1.

Item	Order of patterns		Heard as more extremely stressed		Hypothesis confirmed?	p <
			First patttern	Second pattern		
1.	0	−1	6	21	yes	.01
2.	−2	−1	25	2	yes	.01
3.	−3	−2	23	4	yes	.01
4.	0	+1	4	23	yes	.01
5.	+1	+2	3	24	yes	.01
6.	+3	+2	24	3	yes	.01
7.	0	+1	5	22	yes	.01
8.	−1	0	26	1	yes	.01

The distribution of choices, when tested against a chi-square model (e.g. Guilford 1942: 226ff), was significantly different from random ($p < .01$) for each of the items.

The rationale behind this set-up was that it would not be necessary to compare e.g. pairs like (+3, +1), if the difference between (+3) and (+2) on the one hand, and between (+2) and (+l) on the other, could be detected without problems. Comparisons between such non-contingent patterns were consequently left out. Thus I conclude that the subjects − as a group − were eminently able to tell all the patterns apart.

5.2 Pretest 2

5.2.0 Introduction

In the previous experiment I had my subjects tell the differences among the seven synthesized stress patterns by making two-by-two comparisons. In this experiment I wanted to go one step further, and see if the subjects could order the patterns on a less comparative, i.e. more absolute, basis. To this effect I devised a ‘stress balance’-scale. This device presupposes comparisons of the amount of stress on the vowels of two syllable words. A point in the exact centre of the scale, let us call it ‘0’, symbolizes the situation in which there are no differences in stress between the two vowels, i.e. cases of double stress.

Stronger stress on the first syllable, i.e. a falling stress pattern, is symbolized by a point further to the right on the scale, customarily represented by a positive value. The stronger the falling stress, the further to the right we go on the scale, and the higher the positive value is. Rising patterns, on the other hand, are symbolized by a point to the left of the 0, and hence by negative values.

In the present experiment the subjects were required to place each item they heard on one of a number of fixed positions along this stress balance dimension. Naturally this task is much more difficult than the comparison job described in § 5.1.

5.2.1 Stimuli

The stimulus material was constructed on the basis of the same seven synthesized stress patterns as under § 5.1.1.2.

The test contained 14 items. viz. twice the set of seven patterns in a random order. Each item was preceded by a number and contained one pattern plus an immediate repetition. The items followed at 10-sec. intervals

5.2.2 Subjects

The same group of subjects was used as under § 5.1.2.

5.2.3 Procedure

The subjects received written instructions (for a specimen see Appendix V), in which the concept of the stress balance scale was explained. They were required to listen to the tape, and to encircle the position on their answer sheets (Appendix VI) that in their opinion corresponded optimally to the stress pattern they perceived. The rest of the procedure was identical to the one described under § 5.1.3.

5.2.4 Results and analysis

The results are given in Table 9. Part of that information is represented graphically in Figure 9.

Tables 9 & 10: Results pretest 2. Distribution index (see text) is given in the rightmost column. All distributions differ significantly from chance by a chi-square test (p < .01). Mean = Index/20.

Pattern	Item nr.	Identified as pattern							Index	Mean
		(−3)	(−2)	(−1)	(0)	(+l)	(+2)	(+3)		
(−3)	6	**12**	11	1	2	0	0	1	−56	−2.8
(−3)	13	**11**	12	2	0	2	0	0	−57	−2.9
(−2)	1	1	**6**	14	1	2	3	0	−21	−1.1
(−2)	8	2	**11**	3	3	7	1	0	−22	−1.1
(−1)	4	0	4	**12**	5	4	2	0	−8	−.4
(−1)	12	0	3	**15**	4	4	1	0	−15	−.8
(0)	5	0	1	7	**14**	5	0	0	−4	−.2
(0)	14	0	0	6	**16**	5	0	0	−1	−.1
(+1)	2	0	0	3	11	**9**	4	0	+14	.7
(+1)	11	0	1	4	14	**8**	0	0	+2	.1
(+2)	3	0	1	2	8	11	**5**	0	+17	.9
(+2)	9	0	1	0	5	13	**8**	0	+27	1.4
(+3)	10	0	0	0	3	10	8	**6**	+38	2.6
(+3)	7	1	0	0	0	5	14	**7**	+51	1.9

For none of the 2 × 7 items do we find a random distribution of choices, which is what would have happened had the subjects' reactions been merely a matter of guesswork. The closer the histograms in Figure 9 are to a normal distribution the loss is the perceived stress difference between the two syllables in the crucial words in the items. Negative and positive skew correspond to rising and falling stress patterns, respectively.

As it rather difficult to compare complete distributions for differences, I suggest that it is a better policy to assign an index to the distributions and compare these. The index I have in mind is obtained by weighting the votes in the (−3) and (+3) class with a factor 3, the (−2) and (+2) class with 2, and so forth, and finally adding up the weighted frequencies. Thus the distribution for item 1 gets its index as follows:

1	vote(s) for class	(−3)	−3	×	1	=	− 3
6		(−2)	−2	×	6	=	−12
14		(−1)	−1	×	14	=	−14
1		(0)	0	×	1	=	0
2		(+1)	+1	×	2	=	2
3		(+2)	+2	×	3	=	6
0		(+3)	+3	×	0	=	0
Index							21

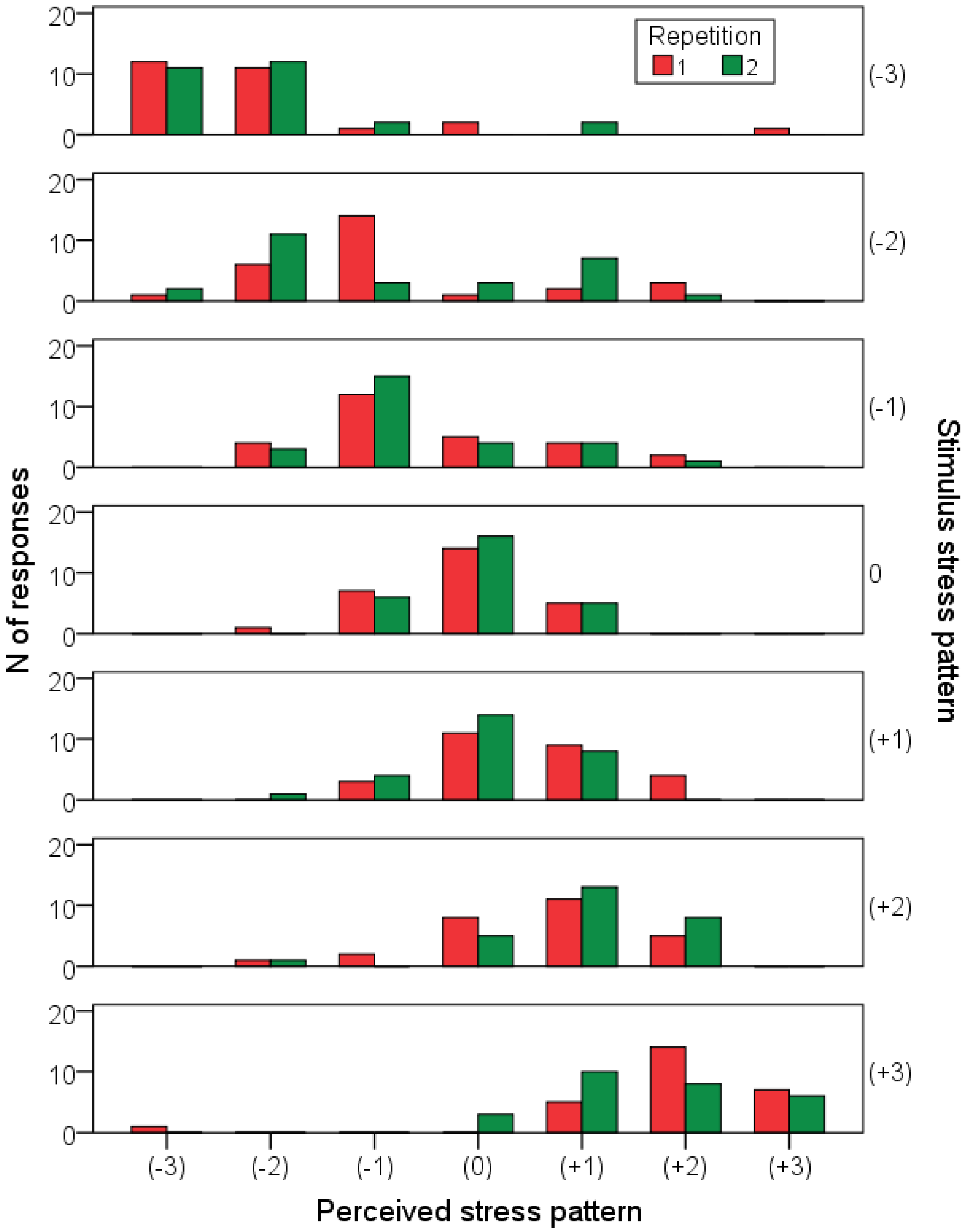


Figure 9: Distribution of responses for pretest 2. First and second presentation of stimuli are kept separate.

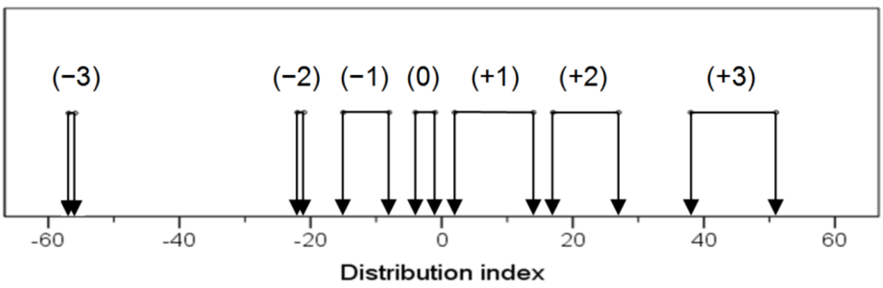


Figure 10: Index and Mean for distributions of pretest 2.

The indices for the distributions are given in Table 9/10. Figure 10, which contains the same information, reveals that:

(1) the indices for the two presentations of the same pattern are never separated by the occurrence of an index of one of the other patterns,
(2) the average indices for each pattern are distributed almost symmetrically round the mid-point of the range, and
(3) the index for the double stress pattern practically coincides with the midpoint of the range.

5.2.5 Conclusions

The conclusions that must be drawn from the two pretests are that the acoustic variations in the stimuli were adequate in so far as they were interpreted by the subjects as variations in stress pattern. Secondly, it follows from the data that the subjects were able to keep the seven stress patterns apart both by comparative and more or less absolute standards.

5.3 Main experiment

5.3.0 Introduction

Research reported on the previous sections has shown that the notion of stress pattern can be operationalized as a single dimension of stress balance in two-syllable words.

In this experiment 1 have artificially created a number of stress patterns corresponding to practically minimally distant points along this stress balance dimension. These patterns are synthesized on a number of words of which we have every reason to believe that their typical stress patterns lie somewhere between the patterns that are characterized by the extreme positive and negative points on the stress balance scale. The thing to find out in each case is which pattern, or group of similar patterns, is chosen by native speakers when asked to select a pattern to match their internal representation of that word in the context given.

With this technique we can investigate a number of problems concerning double stress and rhythmic variations.

In the opening chapters the question of double stress was limited down to the following: is there or is not there a difference between the (1-1) and (2-1) pattern, and if not, do they indeed converge in a (2-1) pattern. The rhythmic principle provided the strategy for testing this: it states that double stressed words are realized as (2-1), (1-1), and (1-2) patterns depending on their phonological context.

Thus it is expected that native speakers will choose differently from among the seven artificial stress patterns depending on the type of word and its context. This will be either a two-way difference, in which case 1 will accept that there is no double stress, or a three-way difference, which would confirm the psychological reality of double stress.

5.3.1 Stimuli

5.3.1.1 Choice of material

In order to keep the amount of material to be synthesized manageable it was decided to incorporate only one double stressed word – with three contexts to elicit (1-1), (2-1), and (1-2) patterns – one instance of rising stress (3-1), and one of falling stress (1-2) in the stimuli.

With this in mind the adjectives *unknown* (1-1), *lifelike* (1-2), and *absurd* (3-1) were substituted for the nonsense word *sisis* in the carrier sentence *Are they .... in 'your country?* (cf. § 5.1.1). Two more sentences were created by inserting the word *quite* immediately before *unknown*, and the word *paintings* immediately after it, satisfying the conditions for a 0_1 and 1_0 context.

5.3.1.2 Synthesis

The synthesis of this material was executed according to the procedure described in § 5.1.1.2. The carrier sentence was kept exactly the same, and the new words *unknown*, *quite*, *paintings*, *lifelike*, and *absurd* were synthesized on the basis of the computer program. No alterations were made to the standard parameter values, except to the durations of the explosion phase of the /k/, and to the subsequent /w/ at the beginning of the word *quite*, reducing the durations to halves.

The necessary F0 information for *quite* and *paintings* was again obtained from narrow-band spectrogram tracings of the same utterance spoken by a native R.P. speaker.

Of each sentence seven versions were synthesized by varying the F0 and A0 parameters for the vowels in the inserted adjectives in exactly the same way as for *sisis* in § 5.1.1.2. For details on the synthesis parameters see Figure 8.

5.3.1.3 Inherent sonority

The crucial words introduced in the carrier sentence do not contain such two identical vowels as in *sisis*. Though the larynx amplitude variations in the seven versions were the same as for *sisis*, the overall intensity differed per vowel, depending on its formant structure. The formant frequency filters of PAT are constructed so as to amplify the harmonics of the larynx pulse in approximately the same way as the human vocal tract does.

Listeners, as Lehiste and Peterson (1959) claim, when asked to determine effort or stress, seem to apply intuitively understood correction factors which subtract the intensity component due to inherent sonority from the overall intensity. It is for this reason, that no steps have been undertaken here to eliminate the artefacts introduced by the formant frequency filters in PAT.

5.3.1.4 Further preparation of stimuli

Each utterance in the synthesized material was recorded on an individual language master card. The recording procedure involved a considerable loss in sound quality, and there is no way of telling to what extent this may have influenced the final results. A normal orthographic representation of the utterance was written on each card, the crucial word underlined. Finally, each card was given a code name, one of a number of common English Christian names.

5.3.2 Subjects

The same group of subjects was used as in the two pretests (§§ 5.1.2, 5.2.2).

5.3.3 Procedure

The subjects took the test individually. They were handed the versions of the first sentence, and asked to play these cards on the language master. It was made clear to them that no two versions were exactly the same, and that the differences were in the stressing of the underlined words in the sentence, as they were written down on the cards. They were then orally instructed to play the cards on the language master, making two-by-two comparisons, and to select the prosodically most natural and satisfactory version from among the seven. Subjects were asked to come up with only one solution, and to exclude the possibility of two equally acceptable versions as a final answer. When the selection of the first sentence was completed they were to write down the code name of their choice on their answer sheets, and to go on to the next set of seven, until all had been finished.

5.3.4 Results and analysis

The results of the test are given in tabular and graphical form (Table 11, Figure 11).

Table 11: Results main test.

Target word	Hypoth. stress pattern	N of subjects indicating preference for synthesized pattern							p	Index
		(−3)	(−2)	(−1)	(0)	(+l)	(+2)	(+3)		
unknown	(1-1)	3	4	3	8	5	2	2	= .500	−5
quite unknown	(2-1)	10	3	3	4	3	2	2	= .100	−26
unknown paintings	(1-2)	4	3	12	4	3	1	0	< .001	−25
lifelike	(1-2)	0	0	0	0	4	3	16	< .001	+58
absurd	(3-1)	2	6	4	2	3	8	2	= .300	+3

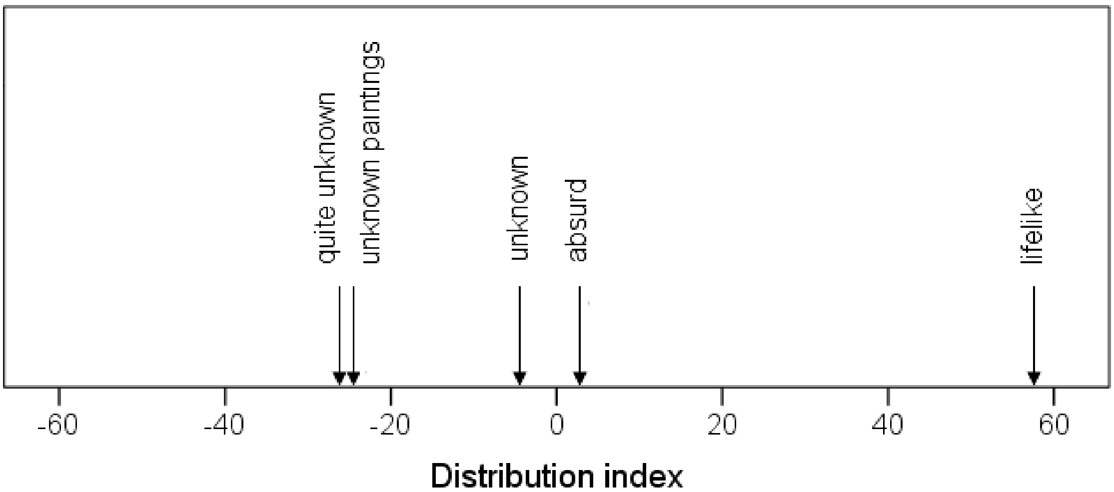


Figure 11-a: Distribution indexes for five target words. The index divided by 20 is identical to the mean of the distribution.

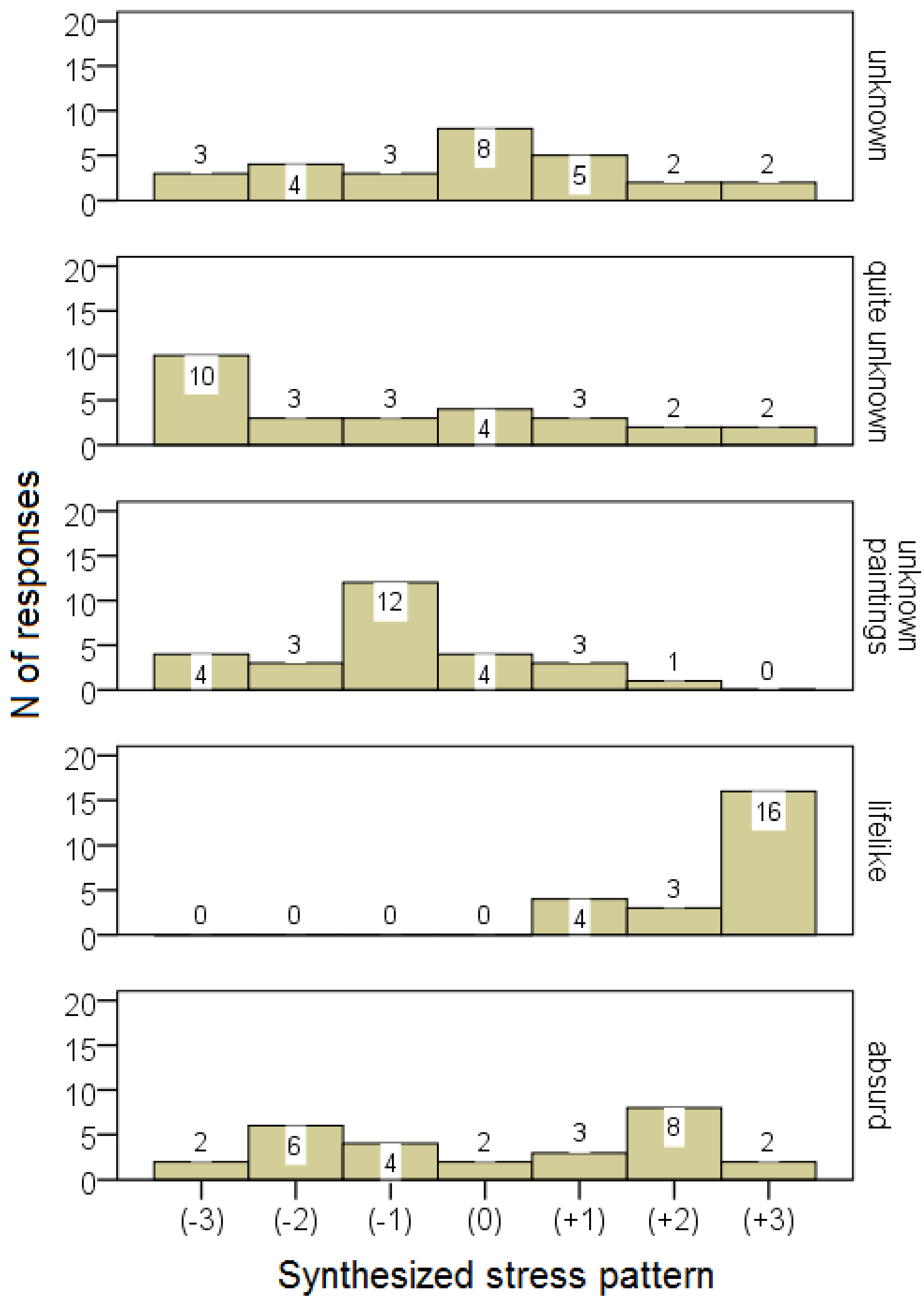


Figure 11-b: Distribution of preferred stress patterns for each of five target words.

As the subjects had been forced to make a choice, the distributions reflecting their preference had to be tested for guessing against a probability model. Chi-square tests (cf. Guilford 1942: 226ff) show that only the distributions for *unknown* in a non-stressed context and *absurd* are non-significantly deviant from chance, which will render interpretation a rather hazardous task.

The distributions were also assigned indices following the procedure outlined in § 5.2.4 (see Table 11). Although this is not reflected in the table, the distributions are all significantly different form each other ($p < .01$ by χ^2).

5.3.5 Conclusions and discussion

(1) The majority of votes for *unknown* in 0_0 context is for the (0)-pattern, and the index for the distribution is −5. Disregarding the statistical insignificance one could say that the native speakers' internal representation of double stress is matched by the level pattern.
(2) In the 1_0 context, where a moderately rising pattern is expected, the subjects' most frequent choice is for the (−3) pattern, with a distribution index of −26.
(3) Comparing the 0_0 and 1_0 contexts we may conclude that there is a clear difference between the double stressed (1-1) and the moderately rising (2-1) patterns. The fact that the two are distinct lends considerable support to the hypothesis that double stress is a separate, psychologically real, category in English. Partial lack of statistical significance denies the results the status of a clear confirmation of the hypothesis.
(4) Counter to the expectation, the results for *unknown* in the 0_1 context do not indicate any preference for a falling pattern. On the contrary, the index is −25, which is about as rising as in the previous case.
(5) In the case of *lifelike*, where the same (1-2) pattern is expected as in (4), we are confronted with more realistic results: an index of +58, and preference for the (+3) pattern. Thus, the rhythmic principle seems to have failed in the 0_1 context.
(6) The results for *absurd* are again counter to our expectations; in fact, the distribution has a +3 index, with a random distribution of votes and, unlike the random distribution for *unknown*, two slight peaks, one in the negative and one in the positive area. Clearly, the subjects must have been completely at a loss with this item. An explanation for this might be the first vowel in *absurd* had its schwa-quality in all versions, i.e. even when stressed. Since reduced vowel quality always coincides with non-stress, all versions must have sounded equally acceptable to the listeners.
(7) The final conclusion must − rather disappointingly − be that the double stress hypothesis is supported by that part of the data that are statistically insignificant, and that the significant results contradict part of the rhythmic principle.

5.3.6 Does double stress exist?

In the analytic part (§ 4.5.2) as well as in the synthesis experiment (§ 5.3.5) we have found some indications for the existence of a separate category of double stress in English. Being very cautious, I should like to put it this way: double stress cannot be refuted on the basis of the results of the experiment reported on in this paper.

Although I believe that the research and analysis techniques that I have developed in the course of this investigation are basically appropriate and insightful, they will have to be refined to a point where statistically significant results become possible. Until then it will not be feasible to settle the matter of double stress in English definitely.

Chapter six

Peripheral experiment 1

6.0 Introduction

This was a very short experiment designed to give me some idea as to the importance of the rhythmic principle from a perceptual point of view.

The rationale behind the experiment is the following. Let us assume that it is true that nominally double stressed words change their (1-1) contour to (1-2) or (2-1) depending on whether a strongly stressed syllable is immediately following or preceding. If we now take an utterance which actually displays such a rhythmic change and cut out the word that conditioned the rhythmic change, eliminate the gap, and play the result to a native speaker, he would get the impression that there is something wrong with the rhythm or stressing of the double stressed word. Should such reactions come up, this would be a strong indication that the rhythmic principle has a certain phonetic reality for native speakers.

6.1 Stimuli

All the utterances of the 1_0 and 0_1 types recorded by informant IV in § 4.2 − as he observed the rhythmic principle more strictly than any of the others − and rerecorded at 38 cm/sec.

By scraping off a few millimetres of magnetically coating on the spot where the conditioning word was, at the time, and playing the result along the playback head of the recorder whenever another few millimetres had been scraped off, the optimal erasure points could be determined. When no trace of the conditioning word could be detected anymore, the scraped off piece of tape was cut out, and the remaining bits spliced together. The original sentences, and the results of the cutting, are given in Table 12. Four out of 14 sentences became ungrammatical after the cutting process; these are starred in the table. The remaining 10 utterances were recorded twice, and each item of an utterance and its repetition was preceded by an identification number. The items followed each other with 10-sec. intervals.

6.2 Subjects

As subjects served 30 native speakers of English, males and females, the majority of them speakers of R.P. English, but supplemented by speakers of other English speaking countries. There was a partial overlap with the group of subjects under 5.1.2, 5.2.2, and 5.3.2, as the number of R.P. speakers available was rather limited.

Table 12: Stimulus material peripheral experiment I. Stimulus sentences were produced by eliminating the stress-shift conditioning context word (indicated by strike-through) from the utterance and splicing the remaining parts together.

1. Things like that are ~~quite~~ unknown in this country
2. Things like that are unknown ~~objects~~ in this country
3. She was ~~just~~ eighteen at the time
4. *There were eighteen ~~girls~~ at the party
5. I want a ~~hot~~ mince pie for dinner
6. *I ate the mince pie ~~hot~~ at dinner yesterday
7. He jumped from the ~~right~~ window on the first floor.
8. He jumped from the window ~~just~~ in time
9. I saw a ~~clear~~ footprint in the garden
10. *There was a footprint ~~right~~ on the spot
11. It is ~~quite~~ absurd to say it
12. *It is an absurd ~~thing~~ to say
13. He'll get the ~~new~~ machine in the morning
14. He'll get the machine ~~back~~ in the morning

Note: Sentences marked with an asterisk were left out of the actual tests. I found these unacceptable, either for grammatical reasons or from the point of view of intonation.

6.3 Procedure

These stimuli were presented to the subjects through an ordinary tape recorder in a normal classroom arrangement.

The subjects were instructed to listen to the utterances on the tape and to pay particular attention to the way in which the crucial words were stressed within the rhythmic structure of the utterance. They had the set of sentences in print on their answer sheets with the crucial words in bold face. They were asked to indicate, by compulsorily encircling one of two options, whether they considered the stress pattern on the crucial word concerned perfectly normal, or in some way odd.

6.4 Results and analysis

The results are summarized in Table 13.

On the assumption that pure guessing in a two-choice situation would have led to a fifty-fifty distribution of normal/odd options, the results were tested for significance against a chi-square model with a probability level of .05. This means that the distribution of preferences must be at least as extreme as 9/21 or 21/9.

Table 13: Results peripheral experiment 1.

Stimulus	Order in presentation	Expected	Response	
			normal	odd
I saw a footprint in the garden	(1)	normal	30	0
It is absurd to say it	(2)	normal	17	13
I ate the mince pie at dinner yesterday	(3)	odd	26	4
I want a mince pie for dinner	(4)	odd	26	4
She was eighteen at the time	(5)	odd	28	2
He jumped from the window on the first floor	(6)	normal	23	7
Things like that are unknown in this country	(7)	odd	25	5
He'll get the machine in the morning	(8)	normal	30	0
Things like that are unknown in this country	(9)	odd	26	4
He'll get the machine in the morning	(10)	normal	29	1

Total score of expectedly normal items: 131 normal versus 19 odd
Total score of potentially odd items: 129 normal versus 21 odd
Difference is insignificant by chi-square test, $\chi^2(2) = .015$.

As we can see at a glance all the items were considered to be natural with unanimity beyond the .05 level, so that the influence of the rhythmic principle seems to be absent: the normal-odd distribution for the normal items taken together is 131/19 versus 129/21 for the sum of the potentially odd items. The difference between these two proportions has a χ^2 of .015, which is totally insignificant.

6.5 Conclusion

Generally speaking, it seems that violation of the rhythmic principle has no dramatic perceptual consequences, a conclusion that might have been drawn earlier, on the basis of the results of the experiment described in § 5.3.

Chapter seven

Peripheral experiment 2

7.0 Introduction

The first aim of this experiment was to investigate if any direct evidence for double stress and rhythmic variations could be obtained from stress judgements on natural utterances. Secondly, it was to provide a measure of perceived stress, which in its turn was to be correlated with a number of acoustic parameters (cf. chapter III). Finally, it was designed to investigate the effects of context on stress judgements. In chapter V the crucial words were presented incomplete utterances. This decision finds its motivation in the results of the present experiment.

7.1 Stimuli

The 28 sentences and 7 words in citation form, as recorded by informant IV − his recording seemed more suitable for the purpose than any of the others − in experiment I, served the basic stimulus material. For the first set of stimuli the crucial words in the 28 complete utterances were isolated from their context. This was done by the cutting and scraping procedure outlined in § 6.1. The cut-out words were recorded twice per item and preceded by a number for identification. The items followed at 10-sec. intervals. The order was the same as in Table 1.

The next set of stimuli were the complete utterances, recorded twice per item and preceded by a number. The 28 pairs of utterances followed at 10-sec. intervals, and their order was as in Table 1. The 7 words pronounced in citation form were then treated in the same way to make up the last set of stimuli.

7.2 Subjects

The same subjects were used as in experiment VI.

7.3 Procedure

The manner of presentation was as under § 6.3. The subjects were asked to encircle that syllable on their answer sheets which they thought was more heavily stressed by the speaker on tape. It was made clear to them that it was a forced choice situation.

After the first set of stimuli new answer sheets were handed out so as to exclude cross-reference. Specimens of instruction and answer sheets are included in Appendix IV.

Table 14: Results of peripheral experiment II.

Item	N of subjects indicating stress on the first syllable when stimulus was presented in		
	context	artificial isolation	citation form
Unknown			
0_0	0	3	3
1_0	0	6	
0_1	20	6	
1_1	4	5	
Eighteen			
0_0	5	14	3
1_0	5	5	
0_1	26	20	
1_1	24	19	
Mince pie			
0_0	3	13	4
1_0	9	6	
0_1	17	24	
1_1	18	14	
Window			
0_0	28	20	30
1_0	29	26	
0_1	27	25	
1_1	25	20	
Footprint			
0_0	27	25	24
1_0	28	29	
0_1	30	28	
1_1	28	28	
Absurd			
0_0	9	5	0
1_0	6	11	
0_1	1	5	
1_1	2	7	
Machine			
0_0	2	16	0
1_0	1	11	
0_1	0	4	
1_1	0	2	

7.4 Results and analysis

We shall refer to the number of subjects that indicate stress on the first syllable of a particular word as the score of that item. Thus the score can never exceed the number of subjects (N = 30). The number of subjects in favour of stress on the second syllable can, of course, be recovered from the score by a simple subtraction. As the information concerning the second syllables is entirely redundant, it will be omitted from the tables and further discussion.

In a forced-choice situation like the one at hand the proportion of votes for each of the two syllables in a word would be fifty-fifty, if there were no audible differences between the two syllables. Therefore, only if the distribution of votes is highly unlikely to have been the result of chance alone, will we accept the view that there must have been an audible stress difference.

As before, I have set a probability level of .05, which means that scores of 9 and less, and 21 and more, are to be interpreted as rising and falling stress respectively; scores less extreme than 9 and 21 are to be interpreted as equal for stress.

The results (i.e. scores) for the three sets of stimuli are given in Table 14.

7.5 Conclusions and discussion

7.5.l Effects of cutting

To get some idea of the general effect of absence versus presence of context on the perception of stress, I plotted the scores for the 28 items in context against those without context. The results can be observed in Figure 12. The unambiguous effect is that presentation in context leads to greater extremity in the scores, whereas cut-out presentation is characterized by greater scatter.

This means that the group of subjects, when considered as a whole, perceived stress more unanimously, more easily, on complete utterances than on artificially isolated words. This was counter to my original belief that subjects would find it easier to give stress judgements on isolated words, where they could fully concentrate on the relevant word, and would not be distracted by the rest of the utterance.

The general effect is captured by the third order regression line in Figure 12, which represents the curve that best fits the coordinates of the X and Y axes.

7.5.2 Absolute double stress

Inspection of the data reveals that none of the citation forms has random scores (i.e. between 9 and 21). This was only to be expected as it appeared in experiment 1 (§§ 4.5.1.2 and 4.5.2) that there are extreme intensity differences in citation forms. It is, however, counter to what is usually claimed in the literature (cf. § 1.4), as we are told there that two equal stresses are preserved in citation forms. Secondly, we find that it is not generally true that random scores come up where double stressed pronunciation is predicted, nor that only extreme scores are found for expected single (i.e. rising or falling) stress.

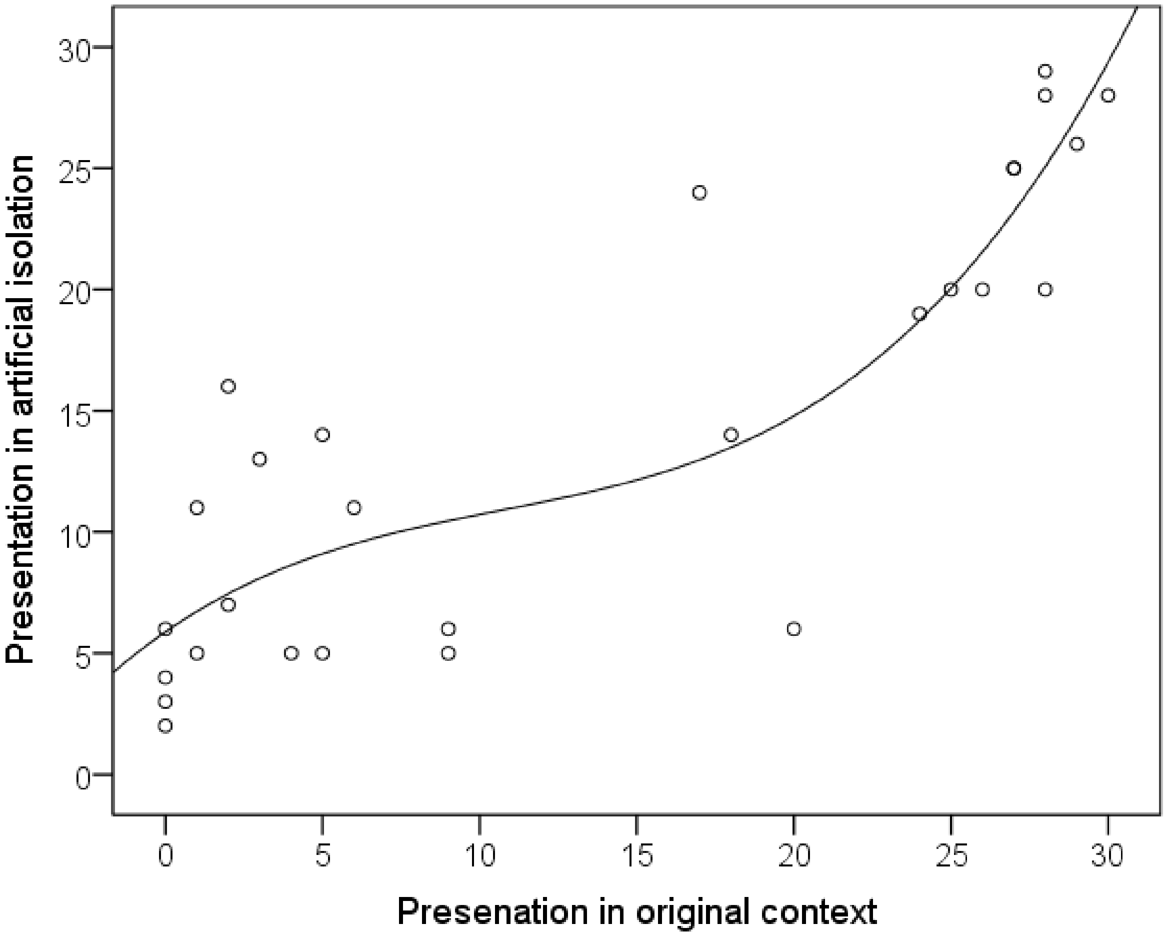


Figure 12: Effect of presentation in context and in artificial isolation.

7.5.3 Relative double stress

We might, however, adopt a weaker interpretation of the data, viz. that the frequency of occurrence of random scores is significantly higher in those cases where double stress is expected than those where we predict failing or rising stress. This new hypothesis has been tested for context and cut-out presentation; once the influence of the rhythmic principle was disregarded, i.e., double stress was predicted for every item involving the words *unknown*, *eighteen* and *mince pie*; the second time only the 0_0 context was considered as potential double stress.

Table 15 gives the data and the values obtained from the association tests performed on them. The effect is non-existing: never do we obtain a significant χ^2 value, so that no evidence for double stress can be abstracted from this experiment

Table 15: Proportions of random versus extreme scores in nominally double stressed versus single stressed words

I. Disregarding the possible influence of the rhythmic principle

Presented in	Nominally double stress		Nominally single stress		Difference between proportions		p
	Random	extreme	Random	extreme			
Context	3	9	0	16	$\chi^2(2) =$	2.24*	ins.
Isolation	5	7	5	11	$\chi^2(2) =$	.32	ins.
Total	8	16	5	27	$\chi^2(2) =$	.46	ins.

* Yates' continuity correction was applied

II: Allowing for the influence of the rhythmic principle.

Presented in	Nominally double stress		Nominally single stress		Difference between proportions		p
	Random	extreme	Random	extreme			
Context	0	3	3	22	$\chi^2(2) =$	.12*	ins.
Isolation	2	1	6	19	$\chi^2(2) =$	.75*	ins.
Total	2	4	9	41	$\chi^2(2) =$	.12*	ins.

* Yates' continuity correction was applied

7.5.4 The rhythmic principle

The influence of the rhythmic principle is quite noticeable in the results of this experiment. The scores on the class of invariably rising stresses centre in the extreme low region with very little variability; the nominally falling stresses are typically found in the high extremes, and are tightly clustered. The double stressed words, however, have an average score of about 15 and an enormous scatter, which, of course, is brought about by the fact that these words are realized as falling stresses on one occasion and as rising stresses on another.

The relevant scores, averages, and variability measures are given in Table 16. The variability measure is the sum of the squared deviations of the individual scores from the average score, divided by the number of items of the type of word concerned.

Table 16: Means and variances of scores

Item	Mean	Variance	N
unknown	5.22	31.65	9
eighteen	14.55	75.24	9
mince pie	12.00	44.44	9
window	25.55	11.35	9
footprint	27.44	3.13	9
absurd	5.11	11.90	9
machine	4.00	2.58	9

Now that we have established that there is no influence of the rhythmic principle in nominally single stressed words (i.e. there is no variability), let us see if the large scatter in the nominally double stressed words can be explained by the rhythmic principle. To this effect we shall consider the scores for nominally double stressed words under 1_0, 0_1, and 0_0 contexts. When the nine items are taken together, the predictions of the rhythmic principle are confirmed 7 times. The data are presented in Table 17.

Generally speaking, the effects of the rhythmic principle are reflected in the data of this experiment.

Table 17: Effects of the rhythmic principle.

Context	Predicted score	Observed score for unknown	Observed score for eighteen	Observed score For mince pie
1_0	under 15	0	5	9
0_1	over 15	20	26	17
1_1	under 15	4	24*	18*

*Counter to prediction

Chapter eight

Peripheral experiment 3

8.0 Introduction

Technically, this was not an experiment at all, as there was no hypothesis involved. It was in fact just an extension of the analysis of the data provided by chapters IV and VII.

In chapter IV it appeared that the duration parameter did not relate to stress, whereas a quite reasonable indication of the presence of stress was given by F0 interval indices, especially when in cooperation with intensity differences. These acoustic measures were taken as a basis for stress patterns as they were predicted in the handbooks. In this experiment I have related the physical measures directly to perceived stress.

8.1 Rank orderings

8.1.1 Ranks for perceived stress

The scores for the three sets of stimuli in §§ 7.1 and 7.3 were ranked separately along the dimension of agreement among the subjects, i.e., the items that had received unanimous stress judgements for the first syllable, having a score of 30, appeared at the top of the rank order; those with unanimous judgements for stress on the second syllable came at the bottom, having a score of 0.

Apart from these rankings, ranks were drawn up for the various realizations of each different word; once the citation forms were included, giving 7 ranks of 5 items; once these were left out, so that there were 7 ranks of 4 items.

8.1.2 Ranks for intensity differences

The intensity differences of the 35 crucial words as pronounced by informant I in § 4.1 (Table 2) were ranked once for all 35 together, once separately for the 28 words in sentences and the citation forms.

In order to avoid tied ranks, of which there would have been many if the differences were based on integral decibel intensity values, I estimated the intensity values again, converting millimetres to 41 dB steps on the intensity calibration graph in Figure 1. A similar set of ranks was constructed for intensity differences when corrected for inherent sonority. The same correction factors were used as in § 4.4.2.3.2.

8.1.3 Ranks for F0-intervals

In a similar fashion the F0 intervals were ranked, the greatest positive intervals at the top, the greatest negative ones at the bottom.

8.1.4 Ranks for duration proportions

For the duration parameter separate ranks were drawn up for the contextually different realizations of each of the 7 words. Here the realization with the largest proportion for the first syllable appeared at the top, the progressively smaller proportions coming further down.

8.2 Correlations

Rank-difference correlation coefficients (cf. Guilford 1942: 305ff) were calculated for F0 intervals, intensity differences, duration proportions with unanimity of stress perception under presentation with and without context and in citation forms. The coefficients are given in Table 18. Intercorrelations were calculated for corrected and uncorrected intensity differences, and F0-intervals. These coefficients are given in Table 19.

I finally calculated the multiple correlation coefficients of each combination of two physical parameters and stress judgement. Note that the multiple correlation coefficients were calculated as if the individual coefficients were based on the product-moment formula, which of course they were not. Hence the multiple correlation coefficients are not accurate, and should only be taken as a useful indication of the factors involved in stress perception. The multiple correlation coefficients are given in Table 20.

8.3 Analysis

It appears that there is no, or even a slightly negative, correlation between the duration proportions and unanimity in stress perception. There is a modest correlation between intensity differences and stress, and even a considerable correlation between F0 intervals and stress perception.

On the basis of these findings the duration parameters will not be used in further analyses. Of the four remaining physical measures, i.e. F0 intervals corrected and uncorrected intensity differences, only the latter two have high intercorrelations. This explains why there are no significant differences between the correlations of corrected intensity and stress perception, and uncorrected intensity differences and stress perception.

The multiple correlations between intensity differences + F0 intervals and stress perception are rather high. It is interesting to note that the intercorrelation is about .1 higher for contexted presentation than for cut-out presentation (cf. § 7.4.1).

8.4 Conclusion and discussion

The general conclusions of this analysis are remarkably similar to those of central experiment I. In both cases we find that a very reasonable prediction of stress can be made on the basis of combined F0 and intensity information, the influence of the correction factor does not seem to be too important, and there are no indications that the duration parameter is involved in ay significant way.

These results do not entirely coincide with the now classical experiments by Fry (1955, 1958). He found that duration variations were very effective as cues for stress perception, that intensity differences were only a bit less effective, and that fundamental frequency had an all-or-nothing effect. It must be clear that the effect of my F0 interval parameter is a gradual one.

Finally, the results of this analysis lend extra support to my earlier decision to vary F0 and intensity, but keep duration constant in the synthesis experiments.

Table 18: Correlation coefficients

A. Duration proportions × stress perception.

Word	Excerpted from spoken context			Presentation in context, citation form,		
	rho	N	sign.	rho	N	sign.
unknown	−.350	4	---	.075	5	---
eighteen	−.800	4	---	.175	5	---
mince pie	−.400	4	---	−.200	5	---
window	.150	4	---	.500	5	---
footprint	.650	4	---	−.125	5	---
absurd	.650	4	---	−.525	5	---
machine	−.300	4	---	.500	5	---

B. Corrected and uncorrected intensity, and F0 intervals × stress perception.

Presentation	Intensity difference						F0 interval index		
	uncorrected			corrected					
	rho	N	p ≤	rho	N	p ≤	rho	N	p ≤
Excerpted	.413	28	.05	.392	28	.05	.673	28	.01
In context	.311	28	ins.	.397	28	.05	.736	28	.01
Citation form	.857	7	.05	.829	7	.05	.598	7	ins.

Table 19: Intercorrelations between acoustic predictors of perceived stress.

Variables	rho	N	p ≤
Uncorrected intensity × corrected intensity	.888	35	.01
Uncorrected intensity × F0 interval index	.249	35	ins.
Corrected intensity × F0 interval index	.332	35	ins.

Table 20: Multiple correlation coefficients *R* obtained for pairs of acoustic predictors of perceived stress

	Presentation type		
	excerpted	in original context	citation form
Uncorrected intensity + F0 interval index	.719	.818	.944
Corrected intensity + F0 interval index	.697	.754	.897

References

Chomsky, N. and M. Halle (1968). *The Sound Pattern of English*, Harper and Row, New York, Evanston, and London.

Essen, O. von (1966). *Algemeine und angewandte Phonetik*. Akademieverlag, Berlin.

Fant, G. (1958). *Modern Instruments and Methods for Acoustic Studies of Speech*, Oslo University Press.

Fourcin, A. and E. Aberton (1971). First Applications of a New Laryngograph, *Medical and Biological Illustration* 21: 172-182.

Fry, D. (1955). Duration and Intensity as physical correlates of English stress. *Journal of the Acoustical Society of America* 27: 765-768.

Fry, D. (1958). Experiments in the Perception of Stress, *Language and Speech* 1: 126-156.

Fuhrken, G. (1932). *Standard English Speech*, Cambridge University Press.

Gimson, A. (1962). *An Introduction to the Pronunciation of English*, Arnold, London.

Guilford, J. (1942 [1965]). *Fundamental Statistics in Psychology and Education*, McGraw-Hill, New York.

Halle M. and S. Keyser (1971). *English Stress. its Form, its Growth and its Role in Verse,* Harper and Row, New York, Evanston, and London.

Helmholtz, H. (1954): *On the Sensations of Tone: a physiological Basis for the Theory of Music*, Dover Publications, New York.

Heuven, V. van (1972). An Experiment on the Relevance of Linguistic Expectancy for the Perception of Stress in Natural Language, Utrecht University, unpublished mimeograph, Dept. English, Utrecht University.

Heuven, V. van (1973). Double Stress and Rhythmic Variation in English Transformational Phonology, unpublished MA thesis, Dept. General Linguistics, Utrecht University.

Holmes, J., I. Mattingly and J. Shearme (1964). Speech Synthesis by Rule, *Language and Speech* 7:127-143.

Jones, D. (1917 [1967]). *Everymen's Pronouncing Dictionary*, Dent, London.

Jones, D. (1918 [1964]). *An Outline of English Phonetics*, Heffer, Cambridge.

Kenyon, J. and T. Knott (1944). *A Pronouncing Dictionary of American English*. Merriam, Springfield.

Kingdon, R. (1958a). *The Groundwork of English Stress*, Longmans, Green and Co., London.

Kingdon, R. (1958b). *The Groundwork of English Intonation*, Longman, Green and Co., London.

Kost H., W. Zinkstok and W. Zonneveld (1972). An Investigation into stress levels in Dutch, Utrecht University, unpublished mimeograph, Dept. English, Utrecht University.

Kruisinga, E. (1918 [1964]). *An Introduction to the Study of English Sounds*, Wolters-Noordhoff, Groningen.

Kurath, H. (1963). A Phonology and Prosody of Modern English, Winter − Universitätsverlag, Heidelberg.

Ladefoged, P. (1971). *Preliminaries to Linguistic Phonetics*, The University of Chicago Press, Chicago and London.

Lawrence, W. (1953). The Synthesis of Speech from Signals which have a Low Information Rate, *Communication Theory*, Butterworth Publications, London, 460-471.

Lehiste, 1. (1970): *Suprasegmentals*, M.I.T. Press, Cambridge Massachusetts and London.

Lehiste, I. and G. Peterson (1959). Vowel Amplitude and Phonetic Stress in American English, *Journal of the Acoustical Society of America* 31: 428-435.

Lehto, L. (1969). *English Stress and its Modification by Intonation. An Analytic and Synthetic Study of Acoustic Parameters*, Suomalainen Tiedeakatemia, Helsinki.

Lieberman, P. (1965): on the Basis of the Perception of Intonation by Linguists, *Word* 21: 40-54.

Lieberman, P. (1967). *Intonation, Perception, and Language*, M.I.T. Press, Cambridge Massachusetts.

McAllister, R. (1971). Predicting Physical Aspects of English Stress, *Speech Transmission Laboratory Quarterly Progress and Status Report* 8-9: 20-29, Royal Institute of Technology, Stockholm.

Morton, J., and W. Jassem (1965). Acoustic Correlates of Stress, *Language and Speech* 8: 159-181.

Netsell, R. (1969). Subglottal and Intraoral Airpressures during the Intervocalic Contrast of /t/ and /d/, *Phonetica* 20: 68-73.

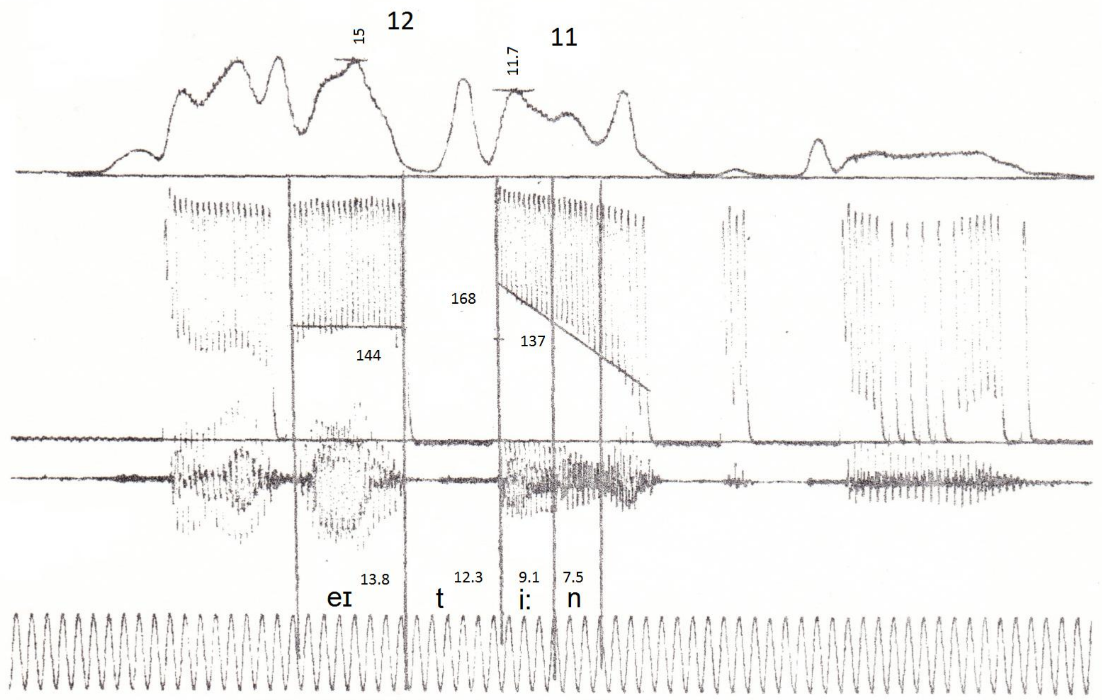
Öhman, S. (1967). Word and Sentence Intonation: a Quantitative Model, *Speech Transmission Laboratory Quarterly Progress and Status Report*, Royal Institute of Technology, Stockholm 2-3: 20-54.

Peterson, G. and I. Lehiste (1960). Duration of Syllable Nuclei in English, *Journal of the Acoustical Society of America* 32: 693-703.

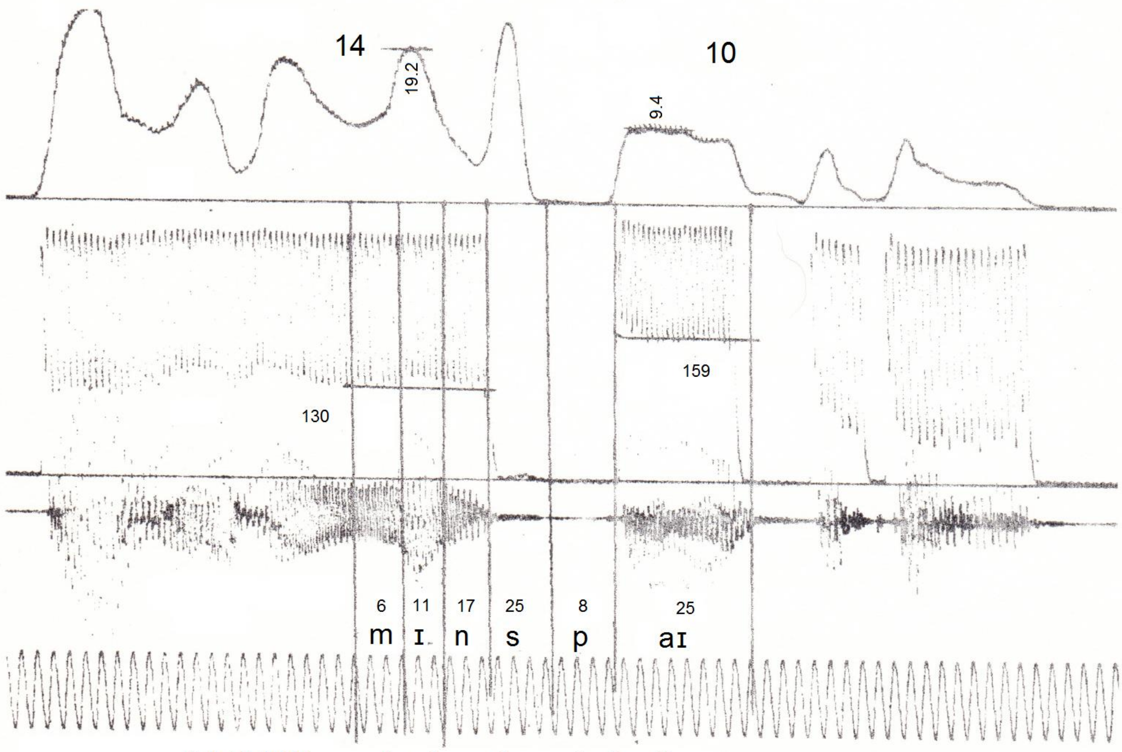
Vanvik, A. (1961). *On Stress in Present-Day English*, Norwegian University Press, Bergen.

Appendix I:

Specimens of mingograms (central experiment I)



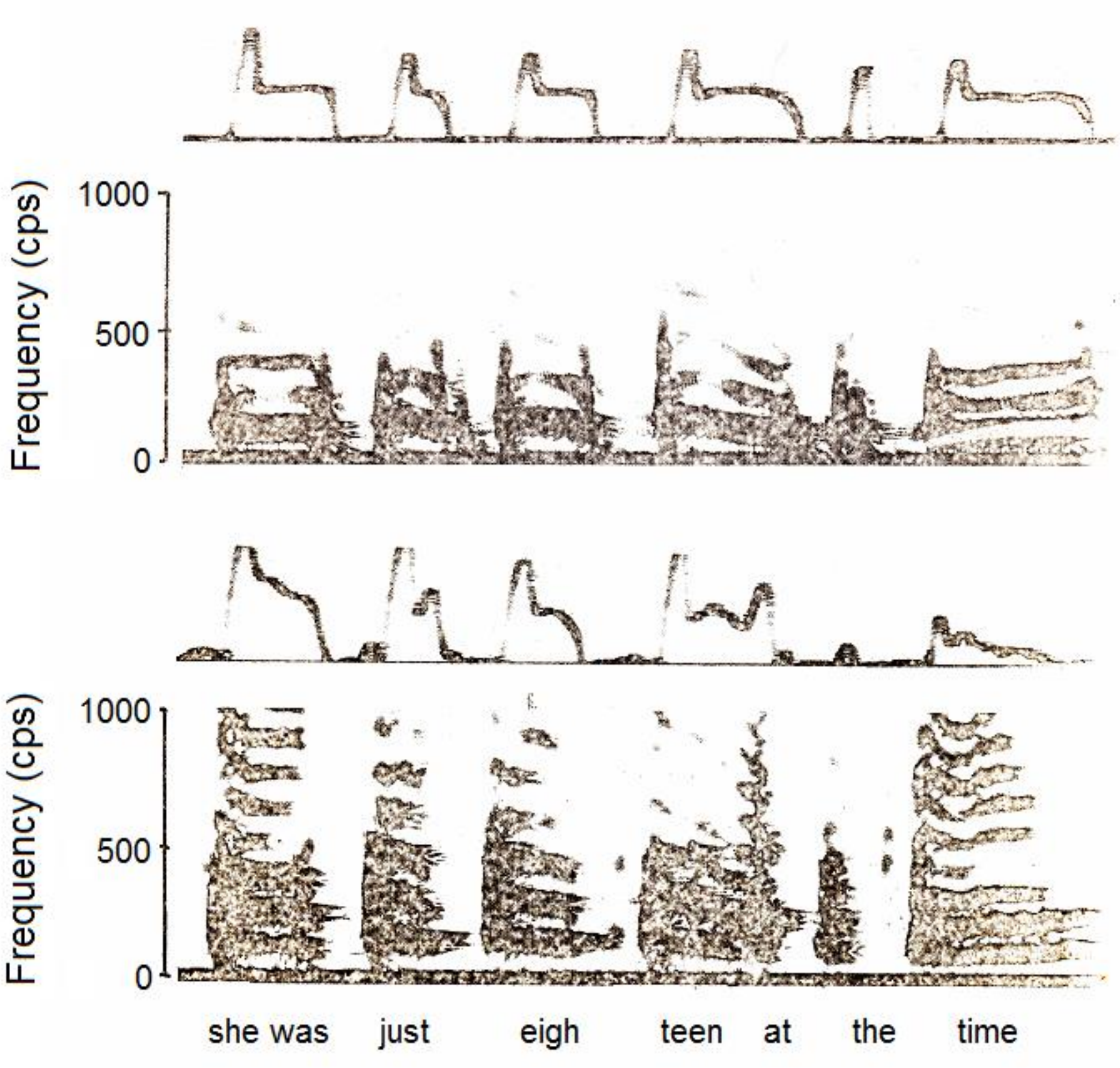
IV (21) She was eighteen at the time



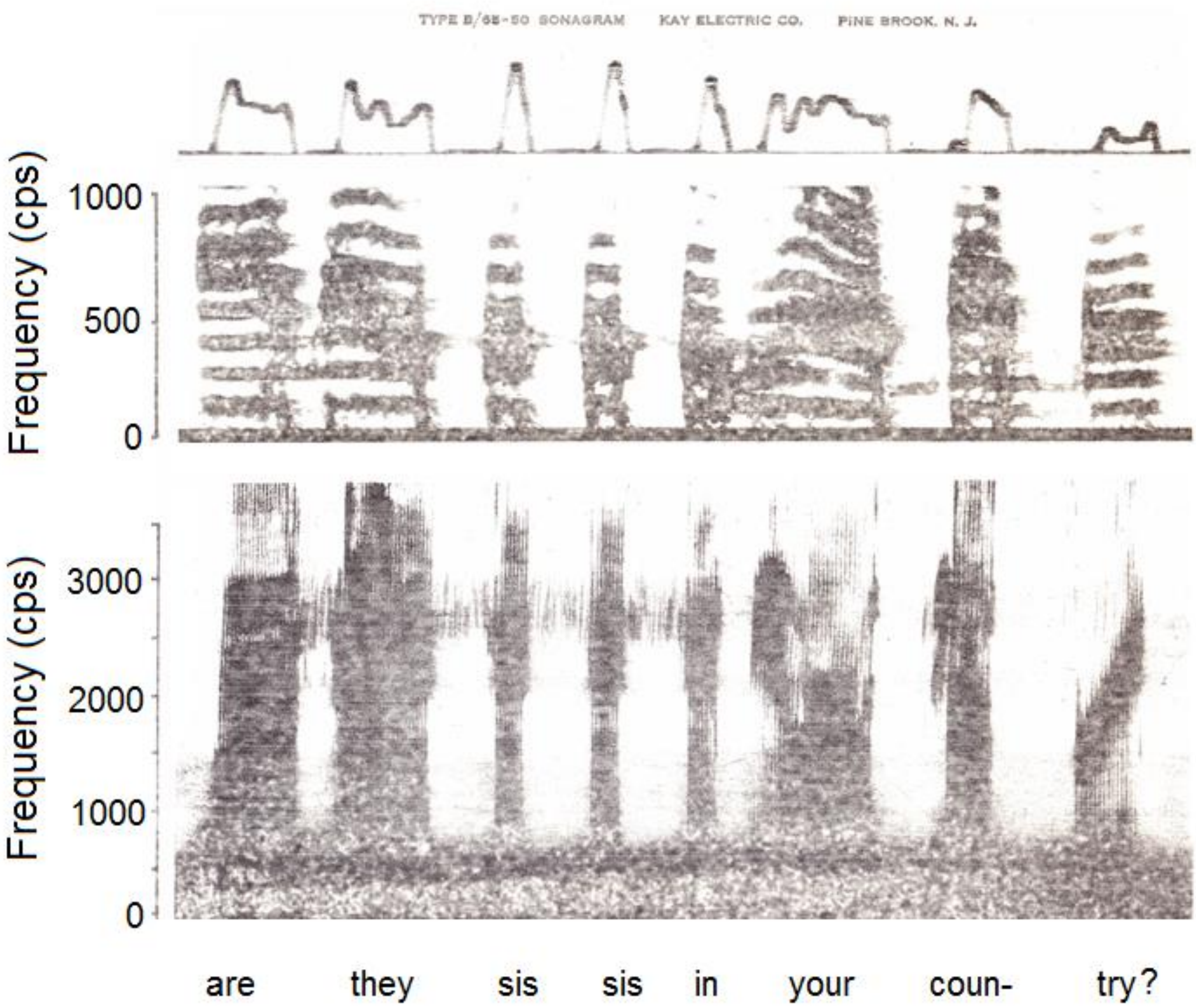
IV (17) We are having mince pie for dinner

Appendix II

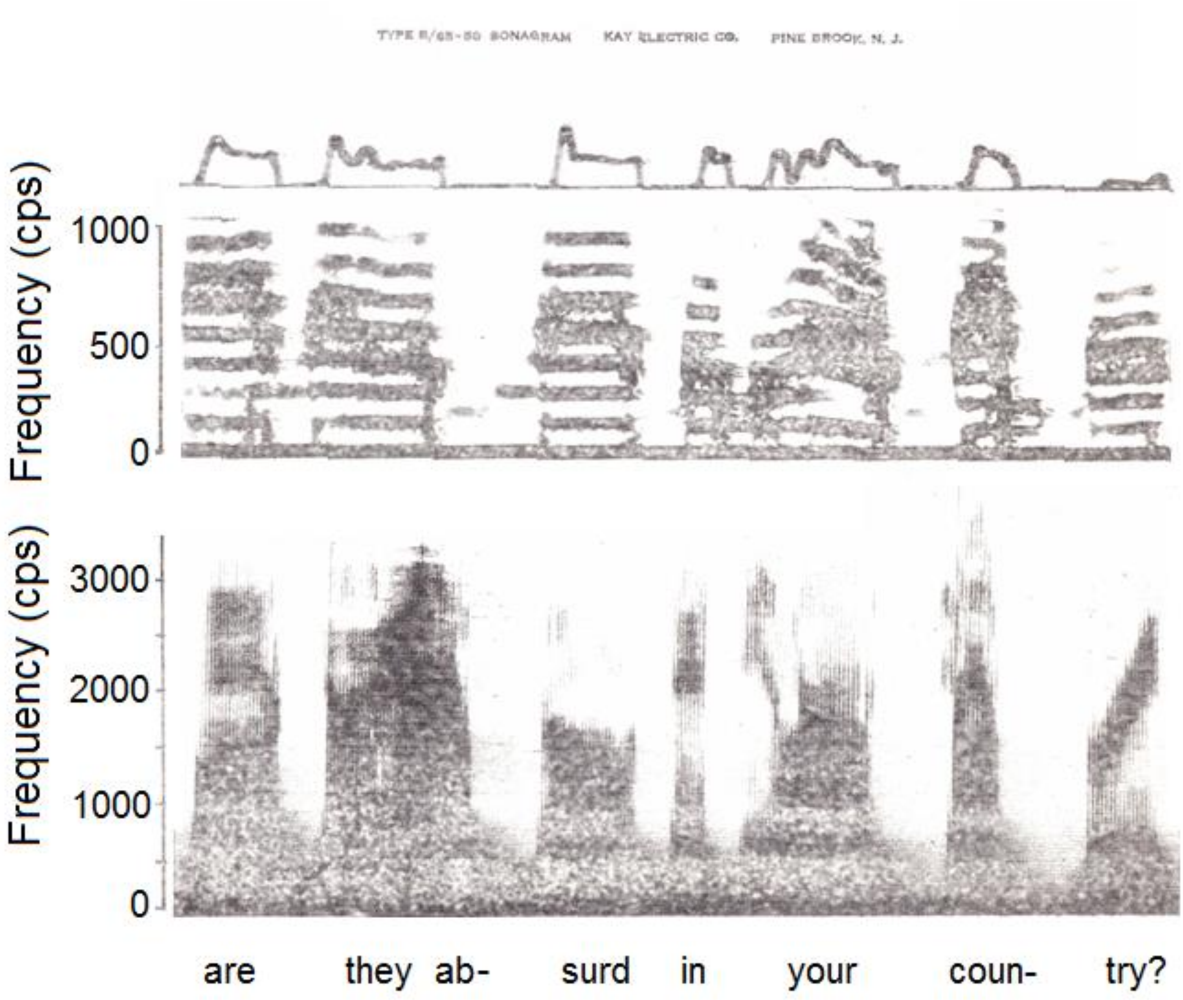
Spectrograms



Intensity trace and narrow band spectrogram of complex natural signal (lower panel) and its glottogaph-pulse counterpart (upper panel) of the same recording (Central experiment I).



Intensity trace, narrow band and wide band spectrogram of one synthesized stimulus (Experiment II, pretest, see §§ 5.1-2).



Intensity trace, narrow band and wide band spectrogram of one synthesized stimulus (Experiment II, main test, see §§ 5.3).

APPENDIX III:

Conversion from synthesis levels to acoustic measures (Chapter V)

Level	F_0 (cps)	F_1 (cps)	F_2 (cps)	F_3 (cps)	FH_2 (cps)	AH_1, AH_2 (dB)
1	52	130	760	1540	1000	0.00
2	55	160	820	1600		3.50
3	59	190	880	1660		5.25
4	63	220	940	1720	1400	7.00
5	67	250	1000	1780		8.75
6	70	280	1060	1840		10.50
7	74	310	1120	1900	1800	12.25
8	77	340	1180	2020		14.00
9	80	370	1240	2080		15.75
10	84	400	1300	2140	2200	17.50
11	90	430	1360	2200		19.25
12	97	460	1420	2260		21.00
13	103	490	1480	2320	2600	22.75
14	109	520	1540	2380		24.50
15	116	550	1600	2440		26.25
16	120	580	1660	2500	3000	28.00
17	126	610	1720	2560		29.75
18	132	640	1780	2620		31.50
19	138	670	1840	2680	3400	33.25
20	146	700	1900	2740		35.00
21	152	730	1960	2800		36.75
22	157	760	2020	2860	3800	38.50
23	164	790	2080	2920		40.25
24	169	820	2140	2980		42.00
25	177	850	2200	3040	4200	43.75
26	188	880	2260	3100		45.50
27	200	910	2320	3160		47.25
28	215	940	2380	3220	4600	49.00
29	226	970	2440	3280		50.75
30	250	1000	2500	3340		52.50
31		1030	2560	1960	5000	54.25

Appendix IV

Instructions + answer sheet chapter 5.1

Introduction

These are some experiments on the perception of English word-stress patterns. The tests involve words consisting of two syllables. I am interested in the relative differences in stress that can be detected between the two syllables in such words.

Instructions experiment I (= pretest II)

You will presently hear a tape containing a number of sentences of the form:
"Are they sisis in YOUR country?"

"Sisis" is a nonsense word, i.e. it does not exist in English. It was used in these sentences because it has two equal syllables, and no real meaning, which makes it easier for you to concentrate on just the stress differences in these words.

Furthermore, the sentences are not spoken by a human speaker, but by a machine. Although the machine can never even approach the human voice, I trust that the sentences will be generally understandable.

You will notice that the stress pattern of the word *sisis* is different from one instance to the next. In some sentences the first syllable is more heavily stressed than the second, in others the reverse will be the case.

I will now play the tape once to give you an idea of the range of the stress differences involved.

play tape

As you will have noticed, the stress differences between the syllables of the word *sisis* are sometimes not even audible (or perhaps they are to you), or they may be very slight, or quite distinct, and so on. Now here is what I would like you to do.

On answer sheet I you will find for each of the fourteen sentences a stress scale, with values ranging between −3 and +3. On this scale 0 means that there is no difference between the stress of the two syllables in the word *sisis*, +1 means that the first syllable is slightly more stressed than the second, +2 that the first syllable is appreciably more stressed than the second, and +3 that the first syllable is much more stressed than the second.

−1, −2, and −3 mean that the second syllable is respectively slightly, appreciably, and very much more heavily stressed than the first.

To make life easier for you the positions on the scale are also identified by graphic representations of the stress differences they stand for:

−3	·●	= second syllable	very much	stronger than the first
−2	○●	= second syllable	appreciably	stronger than the first
−1	◯●	= second syllable	slightly	stronger than the first
0	●●	= first syllable	equal to	the second
+1	●◯	= first syllable	slightly	stronger than the second
+2	●○	= first syllable	appreciably	stronger than the second
+3	●·	= first syllable	cery much	stronger than the second

You will hear each sentence twice, and there is a number spoken before each pair which corresponds to the number on your answer sheets. Immediately after you hear a pair of sentences you have to encircle the position on the stress scale that optimally represents thec stress difference you heard.

NOTE that you must make a choice even if you find it difficult to decide. You may not encircle more than one position on each scale.

ANSWER SHEET TEST I.

Name:
Age:
Sex:
Characterization of dialect:

	·●	○●	◯●	●●	●◯	●○	●·
(1)	−3	−2	−1	0	+1	+2	+3
(2)	−3	−2	−1	0	+1	+2	+3
(3)	−3	−2	−1	0	+1	+2	+3
(4)	−3	−2	−1	0	+1	+2	+3
(5)	−3	−2	−1	0	+1	+2	+3
(6)	−3	−2	−1	0	+1	+2	+3
(7)	−3	−2	−1	0	+1	+2	+3
(8)	−3	−2	−1	0	+1	+2	+3
(9)	−3	−2	−1	0	+1	+2	+3
(10)	−3	−2	−1	0	+1	+2	+3
(11)	−3	−2	−1	0	+1	+2	+3
(12)	−3	−2	−1	0	+1	+2	+3
(13)	−3	−2	−1	0	+1	+2	+3
(14)	−3	−2	−1	0	+1	+2	+3

Appendix V

Instructions + answer sheet Chapter 5.2

Instructions experiment IA (= Pretest I)

In this test you will hear pairs of sentences of exactly the same type as in the previous test. The two members of each pair, however, are different from each other this time, i.e. they are not mere copies of each other. One of the two will display a more extreme (that is better audible) stress difference than the other. Sometimes the more extreme member is the first of the two, on other occasions it is the second.

Your task is to decide for each pair of sentences whether the first or the second member has the more extreme stress difference.

You may indicate your decision on answer sheet IA. If you find that the first sentence of the pair is more extremely stressed, tick the first box (A); if you find the second member more extreme, tick the second one (B). Again, you must make a choice.

To allow you ample opportunity to make up your mind, each pair is given twice; there are pauses of about five seconds after you have heard the two pairs.

ANSWER SHEET TEST IA

	A	B
(1)	☐	☐
(2)	☐	☐
(3)	☐	☐
(4)	☐	☐
(5)	☐	☐
(6)	☐	☐
(7)	☐	☐
(8)	☐	☐

Appendix VII

Instructions and answer sheets Chapters VI and VII

EXPERIMENT A (Chapter VI)

You will hear a number of two-syllable words, each one repeated once, and preceded by an identification number. The pairs succeed each other with intervals of about 10 seconds.

For each word I want you to listen carefully on which of the two syllables the stress falls. If you hear stronger stress on the first syllable, then encircle the corresponding syllable on your answer sheet. If you hear stronger stress on the second syllable, then encircle the second circle on your answer sheet. If you cannot make up your mind, you will have to guess. It is important that you encircle just one syllable per word, no more and no less.

I want to point out that there is no right or wrong in this experiment. We simply need your opinion as to which syllable sound (more) stressed to your ears.

ANSWER SHEET EXPERIMENT A (Chapter VI)

(1)	un	known	(21)	eigh	teen
(2)	ma	chine	(22)	win	dow
(3)	foot	print	(23)	un	known
(4)	ab	surd	(24)	ma	chine
(5)	ma	chine	(25)	win	dow
(6)	foot	print	(26)	un	known
(7)	ab	surd	(27)	ma	chine
(8)	mince	pie	(28)	foot	print
(9)	foot	print	(29)	ma	chine
(10)	ab	surd	(30)	win	dow
(11)	mince	pie	(31)	un	known
(12)	eigh	teen	(32)	foot	print
(13)	ab	surd	(33)	ab	surd
(14)	mince	pie	(34)	eigh	teen
(15)	eigh	teen	(35)	mince	pie
(16)	win	dow			
(17)	mince	pie			
(18)	eigh	teen			
(19)	win	dow			
(20)	un	known			

EXPERIMENT B (Chapter VII)

Our second experiment is a bit more difficult than the first. You will hear ten short sentences, each one repeated once, and each pair identified by a number. The pairs follow with intervals of about 10 seconds. You will see the same sentences printed on your answer sheet. You will see that each sentence has an underlined two-syllable word in it. We want you to pay particular attention to the way in which these underlined words are stressed in the sentences on the tape. You may feel that the stress pattern on some of these words is in some way or other odd or funny. Or, alternatively, you may find that there is nothing unnatural about the word concerned.

If you think the underlined word is somehow odd, encircle the O option on your answer sheet; if you think there is nothing wrong with the underlined word, encircle the N option (O for odd, N for normal).

ANSWER SHEET EXPERIMENT B

(1)	I saw a <u>footprint</u> in the garden	O / N
(2)	It's <u>absurd</u> to say it	O / N
(3)	I ate the <u>mince pie</u> at dinner yesterday	O / N
(4)	I want a <u>mince pie</u> for dinner	O / N
(5)	She was <u>eighteen</u> at the time	O / N
(6)	He jumped from the <u>window</u> on the first floor	O / N
(7)	Things like that are <u>unknown</u> in this country	O / N
(8)	He'll get the <u>machine</u> in the morning	O / N
(9)	Things like that are <u>unknown</u> in this country	O / N
(10)	He'll get the <u>machine</u> in the morning	O / N

EXPERIMENT C (Chapter VI)

Roughly speaking, this experiment is the same as the first. Only this time you will hear the two-syllable words in full sentences – and not in isolation. The instructions are exactly the same as under experiment A.

ANSWER SHEET EXPERIMENT C

(1) Things like that are UN KNOWN in this country
(2) He'll get the new MA CHINE back in the morning
(3) I saw a clear FOOT PRINT in the garden
(4) It's an AB SURD thing to say
(5) He'll get the MA CHINE in the morning

(6) There was a clear FOOT PRINT right on the spot
(7) It's quite AB SURD to say it
(8) I ate the MINCE PIE hot at dinner yesterday
(9) I looked at the FOOT PRINT in the garden
(10) It's a quite AB SURD thing to say
(11) I want a hot MINCE PIE for dinner
(12) There were EIGH TEEN girls at the party
(13) It's rather AB SURD to say it
(14) I'll have a hot MINCE PIE first thing in the morning
(15) She was just EIGH TEEN at the time

(16) He jumped from the WIN DOW just in time
(17) We are having MINCE PIE for dinner
(18) There were just EIGH TEEN girls at the party
(19) He jumped from the right WIN DOW on the first floor
(20) Things like that are UN KNOWN objects in this country

(21) She was EIGH TEEN at the time
(22) He jumped from the right WIN DOW just in time
(23) Things like that are quite UN KNOWN in this country
(24) He'll get the MA CHINE back in the morning
(25) He jumped from the WIN DOW on the first floor

(26) Things like that are quite UN KNOWN objects in this country
(27) He'll get the new MA CHINE in the morning
(28) There was a FOOT PRINT right on the spot