



**CAMBRIDGE STUDIES  
IN LINGUISTICS • 1**

**Prosodic Systems  
and Intonation  
in English**

**DAVID CRYSTAL**

**CAMBRIDGE UNIVERSITY PRESS**

Published by the Syndics of the Cambridge University Press  
The Pitt Building, Trumpington Street, Cambridge CB2 1RP  
Bentley House, 200 Euston Road, London NW1 2DB  
32 East 57th Street, New York, NY 10022, USA  
296 Beaconsfield Parade, Middle Park, Melbourne 3206, Australia

© Cambridge University Press 1969

Library of Congress Catalogue Card Number 69-13792

ISBN 0 521 07387 1 hard covers

ISBN 0 521 29058 9

First published 1969

Reprinted 1972, 1975

First paperback edition 1976

First printed in Great Britain

at the University Printing House, Cambridge

Reprinted in Great Britain by Redwood Burn Limited,  
Trowbridge & Esher

SYNDICS LIBRARY

CAMBRIDGE

# Contents

---

<i>Preface</i>	<i>page vii</i>
1 Some preliminary considerations	1
2 Past work on prosodic features	20
3 Voice-quality and sound attributes in prosodic study	97
4 The prosodic features of English	126
5 The intonation system of English	195
6 The grammar of intonation	253
7 The semantics of intonation	282
<i>Bibliography</i>	<i>309</i>
<i>Index of Persons</i>	<i>361</i>
<i>Index of Subjects</i>	<i>366</i>



# Preface

---

The study of intonation and related features in language is developing more rapidly these days than at any other time. Even a brief look at the extensive bibliography on the subject makes it clear that a great many linguists are concerned that intonation study should not remain 'the Cinderella of the linguistic sciences', as Alan Sharp once put it (1958, p. 151). But despite this activity, there is still precious little description available of intonation and related vocal effects in English or any other language, and there is a marked reluctance to develop any theoretical perspective, or to provide criteria for evaluating different partial descriptions. In this book, then, I have tried to develop just such a theoretical basis for the study of intonation in English, by emphasising the need for intonational phenomena to be integrated within a more general theory of non-segmental phonology. After an introductory discussion of some of the procedural difficulties involved and a review of the history of work on this subject (chapters 1 and 2), I shall outline various fundamental concepts required for the study of sound, relating these to the linguistically relevant notions of voice-quality and prosodic system (chapter 3). Chapter 4 provides a descriptive framework which defines and interrelates all the prosodic contrasts I have found in English, and acts as a theoretical perspective for the detailed study of those pitch contrasts which form the subject-matter of intonation (chapter 5). I discuss the relevance of intonation for grammatical analysis in chapter 6, and I have tried to isolate the main problems hindering progress in the semantics of intonation in chapter 7. In conclusion, I have attempted to give as comprehensive a bibliographical picture as possible, in the hope that this will clarify the 'state of the art' for future researchers. This book should make it clear that there are still a great many questions unanswered and unasked.

In order to present a reasonably comprehensive picture of my approach, I have incorporated into the present book, sometimes in rewritten form, parts of my earlier published work on intonation, and on prosodic features in general. 2.10.2, 4.4 and 5.10.2 contain material

# I *Some preliminary considerations*

---

## **1.1 The need for information**

While there has been a rapid development in the study of English speech during this century, particularly stimulated by the advent of an autonomous and scientifically orientated discipline of Linguistics, it is none the less very plain from looking at the theoretical and practical articles and book headings with regard to the present state of the study that

ficially realised. It is still rare to find a general guide to English speech (as opposed to a specific study of intonation) which pays close and systematic attention to even the majority of variables within the 'non-verbal' or 'suprasegmental' aspects of spoken language, or a grammar which has a section on intonation and stress near the beginning—or anywhere. Again, the absence of any well-defined theory and procedures of analysis has resulted in distortions and vague conceptual terminology in many of the textbooks which purport to be introductions to intonation and related features in English. This is indeed a paradoxical situation for English language pedagogy to find itself in, as it is a phonetic 'residue' of imperfectly learnt prosodic features which is usually the final barrier to the mastery of a foreign language, by maintaining a stubborn accent, on the one hand, and by obscuring the full range of attitudinal contrasts which prosodic contrasts indicate, on the other. Before such a situation can ever hope to be improved, however, it is necessary to evaluate the already available information about these features, and to supplement it by a large-scale, systematic survey of all the variables made use of in English speech. The conclusions presented in this book are based on an analysis of a large sample of English collected for this purpose.

It is understandable that the study of intonation and related features should be in such a state, when one considers the difficulties involved in subjecting this aspect of language to analysis—problems of obtaining reliable information, of defining the range of variables affecting any semantic interpretation, and of identifying and measuring such elusive phenomena as pitch (see 1.3, 2.10.1, 3.7 and 7.2). But by far the most important reason for current inadequacies is a historical one: as will become clear from chapters 2 and 5, the demands of English-language teaching in the early decades of this century produced partial descriptions which, in the absence of sufficient theoretical and descriptive research, regularly involved oversimplification and misinterpretation. Traditions of study were established which have only recently begun to be critically examined, and the extent to which certain aspects of the subject have been neglected is only now being realised. Misleading, impressionistic statements about specific intonation patterns are the most obvious and widespread result of this neglect—a similar state of affairs to that found in English grammar over the past three or four decades (though the enlightenment which has percolated through grammatical description does not seem to have had any effect on intonational study). Such statements usually imply the existence of statistical support (through the use

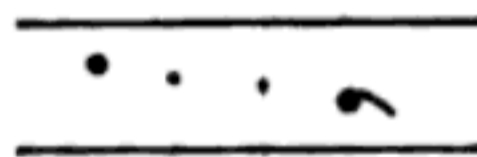


of such terms as 'normally'), but in fact would seem to be based largely or totally on an unscientific impressionism that all too often results in oversimplification of a complex linguistic situation, or in partial truth—which, as is well known, the student-teacher too readily generalises. The observation made by Coleman as long ago as 1914 could be made with equal validity today:

investigations hitherto have been deprived of much of their value through the inquirers' not going far enough afield (thus building up a theory on the few obvious examples that occur to one at the moment) or through their taking their examples from the connected language of books where the sentences are often altogether wanting in the variety found in conversational speech (p. 19).

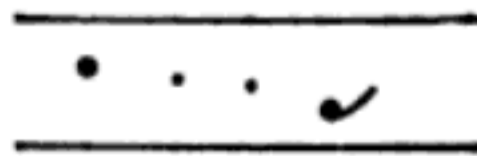
A typical example of the kind of general statement referred to is to be found in most descriptions of the intonation of questions in English. A major distinction is made between the intonation of 'particular' questions (those beginning with an interrogative word such as 'how') and 'general' questions (those arising from inversion of the subject and finite verb): the former are said to be pronounced 'normally' with a falling tone, the latter with a rising tone. Statements of this kind, without any further descriptive qualification (as to what is meant by 'normally', for example) seem to go back at least to Butler (1633), and may be found in almost every exposition of English intonation that has been written.<sup>1</sup> Analysis of most varieties of English speech, however, shows that the issue is hardly as simple as this, it being quite possible to have both a falling and rising tone with each kind of question, the difference lying in the type of attitude involved.

What are you doing?



is generally more serious and abrupt in its implications—at least for British English—than the more friendly and interested

What are you doing?



<sup>1</sup> See, for example, Sweet (1890, p. 32), Palmer (1922, p. 73; 1924, §§439, 444), Armstrong & Ward (1926, pp. 10, 21), Palmer & Blandford (1927, p. 3), Jespersen (1933*b*, §28.6), Bloomfield (1933, pp. 171 ff.), Coustenoble & Armstrong (1934, p. 14), Harris (1944), Allen (1954, p. 43), O'Connor (1954, p. 91), Jones (1956*a*, p. 228), Kingdon (1958*a*, p. 210), and Stockwell (1960*a*). Cf. also Fries (1964, p. 245) for a partial review of this issue.

4 1.1 Some preliminary considerations

and similarly with

Are we going out this evening?  
• . • . ˆ . • .

compared with

Are we going out this evening?  
• . • . ˆ . • .

—all things (for example stress, speed of utterance, facial expression) being equal. It is realistic to generalise only if contextual information is provided and supported by statistics. As Fries says, in the preliminaries to reporting an experiment on this subject (1964, p. 245):

General impressions without the support of any actual counting of the instances in even a small corpus provide no satisfactory ground for generalising concerning comparative frequency of occurrence. Unfortunately, so far as the evidence goes, the many assertions concerning *the rising intonation as the usual mark of yes-no questions in English* have not been based on any adequate body of quantitative information.

Fries's own analysis is stylistically too restricted to provide results of any generalisability (all the questions he examined were used in the context of a quiz-programme), but his findings (that yes-no questions used a falling tone approximately 62 per cent of the time) are not without significance, and an examination of the distribution of tones on such questions in *any* kind of conversational English shows a similar flexibility. Clearly a more complex descriptive statement is suggested, to allow amongst other things for attitudinal and stylistic variables. Occasionally, one does find more satisfactory attempts at description and prescription—Huang & Green (1964), for example—but the enlightenment which Coleman showed (1914, p. 20) in his critique of Jespersen's position in this matter is unfortunately absent from the most widely used textbooks.

The intonation of questions is but one field which has been badly treated; there are many others. It is the presence of such misleading information, plus the absence of any synchronic description of the *full* range of non-segmental vocal effect in English, which provide the two main reasons supporting the need for fresh research into a familiar area. In this book, I hope to be able to carry out two tasks: first, to outline and justify a theoretical framework which will define and interrelate all the non-segmental contrasts which exist in English, and which will allow



used in a number of different senses in early work in this field, some of which are confusing if not vacuous. This can be seen if one examines such common uses as the following, the term being employed differently in each case: 'auditory analysis', 'acoustic analysis', 'analysis of the data' (i.e. to find out what the linguistic contrasts are, presumably), 'analysis' in a statistical sense (cf. p. 10), and 'an analysis' (in the sense of 'structural description'). The latter is the sense in which I am using the term: by 'analysis' I am referring to the explication of the non-segmental contrasts perceived in my data as meaningful (in the sense of p. 19 below) by postulating a set of prosodic systems within which they may be defined and interrelated. I am not referring to the process of recognition which produces our awareness of such contrasts in the first place (so that it is not strictly meaningful to talk of 'analysing' the data, without much qualification); nor am I referring to the different techniques which are available in order to reduce one's data to a form more amenable for analysis (so that I shall not talk of auditory/acoustic/articulatory analysis, but *techniques*).

Secondly, the speech data which provided verification for the analysis were selected to cover a range of educated English, described below, and gathered together as a corpus. The need for such an independent body of data was clear. In a field which deals so closely with personal attitudes, often imprecisely definable by introspection, and difficult to measure and assess, it was not possible or desirable to rely on my own impressions of usage to determine either the form or the function of prosodic features, nor was it possible to determine frequential information in this way. As has been frequently recognised, this is particularly the case when the investigator is a trained phonetician: 'those who have been trained to a conscious control of their intonation patterns cannot provide the body of spontaneous utterances from which to discover exactly what the patterns are and the relative frequency of their use' (Fries, 1964, p. 245). Apart from this, analysis based on the speech-acts of a single person, such as myself, would be unsatisfactory for two reasons: first, I am not certain about my usage of prosodic features in all respects; second, I would find it impossible to say where the boundary line should be drawn between the culturally determined, conventional, linguistically significant features of speech, and the physiologically determined, individual and linguistically uninteresting features (see 3.2, where this distinction is amplified), and my descriptive statements would consequently be 'skewed' by being focused too closely on the idiosyncratic. Reference to information

derived from native speakers other than myself, then, was essential as a check on and a supplement to my own intuition (in no sense is the corpus a *replacement* of this intuition, as is sometimes naïvely suggested), and as a source of suggestions about the internal organisation of language which might have been ignored or overlooked if I had relied solely on introspection.

The actual constitution of the corpus is outlined below (p. 12). Being by definition finite, and incomplete in its attempt to be a representative sample of spoken English, it occasionally needed to be supplemented by information on specific questions. In such cases, the policy used here was to refer to the intuitions of other native speakers (informants), who were as linguistically naïve as possible, by obtaining their reaction about linguistic events in well-controlled experimental situations. Such tests, relating as they did to areas in which the corpus and my own intuition were felt to be inadequate, were not planned in advance as a series of tests, but were rather introduced as the need for them became obvious. When used in the present research, such *ad hoc* testing is referred to as it arises (see 5.3, 7.3). Clearly, if carefully carried out, quite a large amount of useful information about a specific problem can be accumulated fairly rapidly.

The aim of the research, then, was to study the performance (acts of speaking, or utterances) of a number of speakers, gathered together into a corpus, or elicited in informant-reaction tests, as evidence for the definition of the conventional, non-idiosyncratic linguistic system which underlies these utterances. In doing this it is important to stress the consideration, already suggested, that an approach which uses a corpus is 'corpus-based' in only a weak sense of that term: it is not synonymous with 'corpus-restricted'. To go beyond the corpus is envisaged from the very outset, as ultimately one wants to make statements about the language system as a whole—or, to put it in current generative terms, about the underlying 'competence' which linguistic performance is supposed to reflect: 'The problem for the linguist... is to determine from the data of performance the underlying system of rules that has been mastered by the speaker-hearer and that he puts to use in actual performance' (Chomsky, 1965, p. 4; cf. also Chomsky & Halle, 1965, p. 103). To restrict one's attention wholly to the utterances of one's corpus, even though they may show interesting and unexpected patterns, is valuable for such people as the psychiatrist or stylistician, but it is over-limiting for the general linguist, who rather uses this material as a



relatively more objective, public and accessible starting-off point than is otherwise available.

The concentration on performance in many parts of this book is therefore a procedural requirement, and (from the point of view of linguistic theory) a limited aim; but it is none the less a prerequisite in this field, in view of the indeterminacy and intuitive uncertainty about the facts of intonation which the linguist finds. This is an important point, which was only hinted at above. It is not the case that a native speaker has conscious tacit knowledge of *all* the patterns and regularities that constitute English—certainly not in the field of prosodic features (cf. Chomsky, 1965, p. 8). Reliance on intuition is justified until one meets with conflicting intuitions about what is the case, and phenomena like intonation produce queries on every front (though no-one thereby doubts the underlying systemicness of intonation). This in my view is the main justification for any statistical work on language: it should produce information about the facts of English which is unexpected, non-trivial. Of course only such facts will be accepted as linguistically significant as receive *a posteriori* intuitive ratification, otherwise one would let in distributions reflecting all kinds of irrelevant phenomena (cf. the problems of acoustic techniques, p. 13). But the function of intuition here is not as a *basis* of analysis or classification. The statistics display a distribution or pattern which, upon reflection, we realise underlies our unconscious performance. (They may also, of course, confirm our previously held tacit knowledge of patterns, in which case their use would be more trivial; but, as mentioned, our certain tacit knowledge of prosodic features in English is very slight.) In other words, statistics are not brought in at random; they are only introduced at points where the data display no obvious pattern, but reason and intuition tell us that there must be some systemicness present, otherwise the conventional communicative value of prosodic features would break down.

The function of the statistics which lie behind many of the statements in the linguistic description in this book is therefore limited, but essential: it is a supporting and clarificatory role, and not usually an explanatory or evaluative one; it normally verifies (or fails to verify) hypotheses reached on intuitive grounds, but it may sometimes act as a stimulus for the formulation of new hypotheses about language. But the potential relevance of statistics has rarely been given explicit recognition in linguistic research (though cf. Classe, 1939, Herdan, 1966), being but obliquely referred to through the use of inspecific adverbials of frequency,



as mentioned above, or through the introduction of a (undefined) norm of some kind—see, for example, Smits Van Waesberghe (1957, p. 374), Classe (1939, pp. 36–7), or Chreist (1964, p. 45).<sup>1</sup> One must not of course be distracted by the quantitative side of the exercise: non-quantitative information on procedures is required at three separate points—in the preliminary linguistic analysis, in the initial sampling, and of course in the interpretation of the statistical results. The basic point to be remembered by the linguist embarking on research in this field is that statistical analysis must itself be dependent on a certain minimum of prior established qualitative analysis of some kind: obviously one has to have *something* to count at the beginning, to have some structural units defined, and moreover to be fairly sure that what is to be counted is in fact worth counting (cf. Reed, 1949, pp. 235 ff.); and such issues are decided independently of the actual choice and utilisation of statistical techniques.

Problems of sampling are also a separate issue, and are bound up with the nature of the statistical model one is using. The main principle of any such model is that language is viewed as a class of events grouped into categories, each event having a probability of occurrence. This probability is readily definable (cf. Newman, 1957, Connolly & Sluckin, 1962) as the ratio of the number of times an event in a category does occur to the number of times it might have occurred, i.e. to the total number of possible events. It is axiomatic in this approach that 'The most likely value of a proportion in a total set, that is, of the true probability, is the value observed in a limited sample' (Newman, 1957, p. 119). But this is only applicable if it is the case that the initial sampling is adequate, and is as random a sample as possible of the total linguistic population which one is trying to define (see below).

Because of the existence of widely divergent linguistic varieties of spoken English (formal religious monologue alongside informal sports commentary, for instance), it is clearly impossible to achieve any such degree of generality as 'a sample representative of the language as a whole'. Instead of truly random sampling (probably impossible in language study—cf. Newman's scepticism (1957, p. 122)), the linguist must begin by making a controlled sample of utterances, choosing texts for analysis within a relatively restricted and clearly defined range of

<sup>1</sup> This concern for norms is reflected in the comparable optimism of other branches of linguistics, such as stylistics; see Gregory & Spencer (1964, p. 102), Quirk (1961a, pp. 216 ff.).

usage, which is intuitively homogeneous as far as linguistic features *other than* those he is studying are concerned. Thus, before beginning to study prosodic features, one would choose a corpus of utterances as homogeneous as possible from the point of view of regional dialect, register, formality of discourse, and so on in respect of segmental phonetic characteristics, grammar, and vocabulary (for a discussion of the relevant stylistic categories involved, see Crystal & Davy, 1969). The investigator is here concerned simply with minimising the most influential variables on the form of the utterance. He need not be too anxious to obtain absolute consistency in this matter. Apart from the fact that this will be impossible in any case in view of the ever-present idiosyncratic or person-identifying elements in utterance, there is also the point that statistically valid reasoning may be made without such initial total consistency. As Reed says (1949, p. 246): 'To pursue a practical working method, the quantitative linguistic investigator must collect his samples of evidence from *a variety* of sources within the area of the language that he is investigating. In this way, he trusts that those factors which may tend to prejudice the results of his study will cancel one another out' (my italics). Having said this, however, it must be pointed out that within the stylistic and regional limitations of my corpus as it stands (see below), the selection of the different speakers *was* carried out on a random basis; that is, it was judged that there was as much likelihood of utterances of speaker *A* in text *X* occurring in the situation<sup>1</sup> of text *Y*, and so on.

In order to attain a reasonable degree of linguistic homogeneity, then, the speakers whose utterances were analysed were all chosen from a single dialect which I have called, following other uses of the label, educated British English (cf. Jassem, 1949, Dietrich, 1956, Quirk, 1957, 1960, Nyqvist, 1962); the range of accents used would generally be classed together as kinds of Received Pronunciation (see Abercrombie, 1956, Gimson, 1962). This choice was directed, not of course by any concept of a linguistic betterness within this dialect and accent, but by the greater usefulness of research based upon them, as opposed to any other British regional or class dialect or accent: this variety is the one

<sup>1</sup> For the notion of 'situation', see Crystal & Davy (1969, chap. 1). It should also be clear that I am only concerned in this book with the direct representation of prosodic features in the spoken medium: I am not concerned with the reflection of these contrasts in conventional orthography (such as the way in which authors attempt to reproduce the effect of different 'tones of voice' in their work) or in any other medium—these are related but logically later subjects of study.