

Prosody, Phonology and Phonetics

John C. Wakefield

Intonational Morphology

 Springer

Prosody, Phonology and Phonetics

Series Editors

Daniel J. Hirst, CNRS Laboratoire Parole et Langage, Aix-en-Provence, France

Hongwei Ding, School of Foreign Languages, Shanghai Jiao Tong

University, Shanghai, China

Qiuwu Ma, School of Foreign Languages, Tongji University, Shanghai, China

The series will publish studies in the general area of Speech Prosody with a particular (but non-exclusive) focus on the importance of phonetics and phonology in this field. The topic of speech prosody is today a far larger area of research than is often realised. The number of papers on the topic presented at large international conferences such as Interspeech and ICPhS is considerable and regularly increasing. The proposed book series would be the natural place to publish extended versions of papers presented at the Speech Prosody Conferences, in particular the papers presented in Special Sessions at the conference. This could potentially involve the publication of 3 or 4 volumes every two years ensuring a stable future for the book series. If such publications are produced fairly rapidly, they will in turn provide a strong incentive for the organisation of other special sessions at future Speech Prosody conferences.

More information about this series at <http://www.springer.com/series/11951>

John C. Wakefield

Intonational Morphology

 Springer

John C. Wakefield
Department of English Language and Literature
Hong Kong Baptist University
Kowloon Tong, Hong Kong

Department of Asian Studies
Palacký University
Olomouc, Czech Republic

ISSN 2197-8700

ISSN 2197-8719 (electronic)

Prosody, Phonology and Phonetics

ISBN 978-981-15-2263-5

ISBN 978-981-15-2265-9 (eBook)

<https://doi.org/10.1007/978-981-15-2265-9>

© Springer Nature Singapore Pte Ltd. 2020

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

The publisher, the authors, and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

*I dedicate this book to the memory of my father,
Jerry J. Wakefield. Everything good in me
comes from him. The rest is my own doing.*

Acknowledgments

I could never have completed this book without the help of many people at different stages of the project. There are too many to name, but I will mention a few. Robert Bauer was instrumental in getting me into my PhD program at HKPolyU and has always been a great mentor for my research related to Cantonese. Tang Sze-Wing supervised my PhD and has taught me a great deal over the years and still motivates me to work hard to this day. Cliff Goddard, who has always been very approachable and encouraging, helped me with some of my definitions of discourse meanings in Chap. 6. Daniel Hirst motivated me tremendously as a PhD student by expressing interest in my research. He has given me crucial comments and advice about the contents of the book and graciously provided the script that was used for all of the PRAAT figures. Liliane Haegeman, who is another friendly and approachable linguist, has been an inspiration to me and has taught me a great deal. Enoch Aboh has been very encouraging about my work and has given me very helpful comments. Lian-Hee Wee has been the source of many fruitful discussions and was kind enough to read and comment on a draft of the book. Noam Chomsky, who I unfortunately have never met, is responsible for my initial interest in linguistics; I hope he considers this book a worthwhile contribution to the field. I consider all of these people to be mentors and hope they are pleased with the final product of my work. Of course, I am solely responsible for any and all errors that remain.

I thank all those who, over fine drink and food and on many occasions, have demonstrated that it is possible to be an academic and maintain a sense of humor: František Kratochvíl, Jason Polley, Mark Wallbanks, and Nick Williams.

I thank the ambilingual participants, who generously volunteered their time to provide the translations that make up the data in Chap. 6.

And finally, I would never have amounted to much at all without the support of my wife and daughter, Mona and Angie—I thank them most of all.

This research was partially supported by a Hong Kong Baptist University Faculty of Arts research grant: *Discovering Cantonese Semantic Primes and Defining Sentence-final Particles* [FRG2/12–13/060]. The Faculty of Arts also provided use of the HKBU Phonology Laboratory. During my time working as a Key Researcher at Palacký University, Olomouc, this work was also supported by the European Regional Development Fund Sinophone Borderlands—Interaction at the Edges CZ.02.1.01/0.0/0.0/16_019/0000791.

List of Abbreviations

1s	First person singular
2s	Second person singular
3s	Third person singular
-pl	Plural marker
Adv-M	Adverbial marker
ASP	Aspect marker
C ⁰	Head of complementizer phrase
CL	Classifier
CM	Comparative marker
COP	Copula
CP	Complementizer phrase
D	Discourse element/Determiner
DM	Delimitative marker
DP	Determiner phrase
EST	Extended Standard Theory
EXP	Experiential marker
F ₀	Fundamental frequency
FinP	Finite phrase
FocP	Focus phrase
GEN	Genitive marker
L1	First language
L2	Second language
MP	Modal particle
NDP	Null discourse particle
NEG	Negation
NSM	Natural Semantic Metalanguage
P	Proposition
PERF	Perfective marker
PFX	Noun class prefix
PROG	Progressive marker
PRT	Particle

REL	Relative clause marker
SAI	Subject-auxiliary inversion
SM	Subject marker
SFP	Sentence-final particle
SPEC	Specifier
T ⁰	Head of tense phrase
TNS	Tense
TopP	Topic phrase
TP	Tense phrase
VP	Verb phrase
X	Variable to represent element put into focus by “only”
X ⁰	Head of phrase category X
X'	Constituent of category X between the word and phrase levels
XP	Phrase of category X

Contents

1	Introduction	1
	References	2
2	The Forms and Functions of Intonation	5
2.1	The Functions of Suprasegmentals	6
2.2	A Definition of Intonation Based on Its Functions	12
2.3	The Forms of Suprasegmentals	14
2.4	Concluding Remarks	18
	References	18
3	Intonational Meaning	21
3.1	The Nature of Intonational Meaning	21
3.1.1	Context-Dependent Versus Context-Independent Meaning	23
3.1.2	Compositional Versus Holistic Meaning	25
3.1.3	Phonological Similarity and Homophony	26
3.1.4	Gradient Versus Categorically Distinct Forms and Meanings	28
3.1.5	The Linguists Theory of Intonational Meaning	28
3.1.6	Testing the Linguists Theory of Intonational Meaning	29
3.2	Intonation and Discourse Particles	31
3.2.1	Intonation and Segmental Particles Are Two Forms of the Same Thing	33
3.2.2	The Similar Debates About Particle and Intonational Meanings	34
3.3	Concluding Remarks	37
	References	37
4	Evidence of the Morphological Nature of Intonation	41
4.1	Tonal Grammatical Particles and Their Segmental Counterparts	42

4.2	Tonal Discourse Particles and Their Segmental Counterparts	47
4.3	Concluding Remarks	51
	References	52
5	Evidence via Cantonese	55
5.1	The Cantonese Language	55
5.1.1	Why Cantonese Is Ideal for This Kind of Research	56
5.1.2	Intonation in Cantonese	57
5.1.3	Cantonese Sentence-final Particles	58
5.2	The Design of the Research	63
5.2.1	The Participants	65
5.2.2	The Corpus and the Dialogues	66
5.2.3	Data Collection	67
5.2.4	Data Analysis	69
5.3	Defining Sentence-final Particles	73
5.3.1	The Natural Semantic Metalanguage Theory	73
5.3.2	Defining Sentence-final Particles with the Natural Semantic Metalanguage	75
5.4	Concluding Remarks	78
	References	79
6	The Results of the Research	83
6.1	Two Evidential Particles: <i>lo1</i> and <i>aa1maa3</i>	84
6.1.1	The Particle <i>lo1</i>	84
6.1.2	The Particle <i>aa1maa3</i>	107
6.1.3	Summary and Analysis	127
6.2	Two Question Particles	136
6.2.1	The Particle <i>me1</i>	136
6.2.2	The Particle <i>aa4</i>	152
6.2.3	Summary and Analysis	160
6.3	Two “Only” Particles: <i>ze1</i> and <i>zaa3</i>	167
6.3.1	The Particles <i>ze1</i> and <i>zaa3</i>	167
6.3.2	Summary and Analysis	185
6.4	Concluding Remarks	190
	References	191
7	The Syntax of Intonation	195
7.1	Background Information	195
7.1.1	Intonation and Syntax	196
7.1.2	Cartographic Syntax	197
7.2	Tonal Morphemes That Function as Grammatical Particles	198
7.3	Tonal Morphemes That Function as Discourse Particles	201
7.3.1	The Syntax of Polar Interrogative Particles	202

7.3.2 The Syntax of Discourse Particles	213
7.4 Prosodic Structure	223
7.5 Concluding Remarks	227
References	227
8 Conclusions and Implications	231
References	236
Appendix	237
References	243

Chapter 1

Introduction



This book is an updated and extended version of my Ph.D. thesis (Wakefield 2010), and portions of it have been published in various forms elsewhere (Wakefield 2012, 2014, 2016, *in press*). The book's title was inspired by Ladd's (2008) book *Intonational Phonology*. It is widely accepted that intonation is phonological, though there is still much debate about its phonological features; in contrast, there is less consensus about what Ladd (2008: 41) refers to as the "Linguist's Theory of Intonational Meaning, ... [t]he central idea of [which] is that *the elements of intonation have morpheme-like meaning*" (emphasis in italics his). Among those who adopt this view, few take as strong a stance about the morphemic nature of intonation as is proposed here. In the chapters that follow, I argue that intonation's phonological components represent morphemes that exist in the lexicons of speakers' minds, and that these morphemes occupy syntactic positions within the structure of the sentence. This idea is not new; in essence, I am following in the footsteps of Hirst (1977, 1983, 1993), who has made attempts to align intonation with the theory of generative syntax. To some linguists, these will sound like strong, unwarranted views; at the very least, they will wonder what kind of evidence there is to support the hypothesis that intonation is no different from the rest of language, other than its form. The purpose of this book-length treatment of the subject is to clarify precisely what this hypothesis entails and to present evidence in support of it.

Intonation is arguably the most controversial feature of language. It has suprasegmental forms and abstract meanings, making it extremely challenging to analyze and describe. There is much disagreement about its forms and functions, and further complicating the matter, the meanings and uses of many of the technical terms that describe intonation are not consistent throughout the literature. Yet another complicating factor is the widely-held assumption that only some suprasegmentals are part of the linguistic structure; the remaining have been variously described as paralinguistic (Couper-Kuhlen 1986; Ladd 2008), nonlinguistic (Fox 2000), or a form of animal communication (Gussenhoven 2004). These two types of suprasegmentals, i.e., linguistic versus paralinguistic, not only occur simultaneously along with the linear stream of an utterance's segments but also simultaneously with each other. Referring to the two types of supra-

segmentals, Ladd (2008: 6) said, “it is a matter of considerable controversy which aspects are which, or whether such a distinction is even possible.”

Even though it remains unclear how to physically and perceptually isolate linguistic from paralinguistic suprasegmentals, it is assumed throughout this book that the two types are qualitatively different in nature. Linguistic suprasegmentals exhibit an arbitrariness of the sign that is characteristic of other grammatical components within language (Hirst 1983; Couper-Kuhlen 1986; Fox 2000), while paralinguistic suprasegmentals are used to express emotions and attitudes such as fear, anger, impatience, or boredom and are assumed to be fundamentally the same cross-linguistically—though there are of course differences in use and production that stem from sociocultural and individual differences.

This book does not take on the monumental task of teasing apart the physical and/or the perceptual differences between the forms of linguistic and paralinguistic suprasegmentals—something for which other linguists are more qualified than myself. Rather the goal here is to propose a way of conceptualizing the differences between the two, based on the assumption that the two are qualitatively distinct. To this end, I will propose that the term *intonation* be redefined based on what I propose its functions to be.

This book has two goals. The first is to propose how intonation should be conceptualized and recategorized based on the hypothesis that it is morphemic. This is dealt with in Chaps. 2 and 3. Chapter 2 describes the forms and functions of intonation and then offers a definition of it based on the functions I propose it has. Chapter 3 discusses the meanings of discourse intonation and contrasts them with the meanings expressed by segmental discourse particles. The second goal of the book is to present evidence and arguments in support of the hypothesis—both from the literature and from my own research. Chapter 4 reviews the morphemic nature of tones that have been reported in the literature, which includes tonal morphemes that have grammatical functions, as well as those that have discourse meanings. Further evidence of discourse tonal morphemes is presented in Chaps. 5 and 6, where my own research is reviewed, offering empirical evidence to indicate that specifically shaped pitch contours in English have definable, context-independent meaning. Chapter 7 proposes how intonation might be represented in the syntactic structure, which is, as far as I know, the most comprehensive proposal presented to date on the syntax of discourse intonation. Finally, Chap. 8 offers some concluding remarks and discusses the implications of analyzing intonation in this way.

References

- Couper-Kuhlen, E. (1986). *An introduction to English prosody*. London: Edward Arnold.
- Fox, A. (2000). *Prosodic features and prosodic structure: The phonology of suprasegmentals*. Oxford: Oxford University Press.
- Gussenhoven, C. (2004). *The phonology of tone and intonation*. Cambridge: Cambridge University Press.

- Hirst, D. (1977). *Intonative features: A syntactic approach to English intonation*. The Hague: Mouton.
- Hirst, D. (1983). Interpreting intonation: A modular approach. *Journal of Semantics*, 2(2), 171–182.
- Hirst, D. (1993). Detaching intonational phrases from syntactic structure. *Linguistic Inquiry*, 24(4), 781–788.
- Ladd, D. R. (2008). *Intonational phonology* (2nd ed.). Cambridge: Cambridge University Press.
- Wakefield, J. C. (2010). *The English equivalents of Cantonese sentence-final particles: A contrastive analysis*. Unpublished doctoral thesis, The Hong Kong Polytechnic University, Hong Kong.
- Wakefield, J. C. (2012). A floating tone discourse morpheme: The English equivalent of Cantonese 㗎. *Lingua*, 122(14), 1739–1762.
- Wakefield, J. C. (2014). The forms and meanings of English rising declaratives: Insights from Cantonese. *Journal of Chinese Linguistics*, 42(1), 109–149.
- Wakefield, J. C. (2016). Sentence-final particles and intonation: Two forms of the same thing. In J. Barnes, A. Brugos, S. Shattuck-Hufnagel, & N. Veilleux (Eds.), *Speech prosody 2016* (pp. 888–892). Boston: Boston University. <https://doi.org/10.21437/SpeechProsody.2016>.
- Wakefield, J. C. (in press). It's not as bad as you think: An English tone for downplaying. In W. Gu (Ed.), *Studies on tonal aspects of languages*. Hong Kong: Journal of Chinese Linguistics Monograph.

Chapter 2

The Forms and Functions of Intonation



The term *intonation* is restricted here to refer only to prosodic features that are morphemic. The term *prosody* is used more broadly to refer to two types of suprasegmentals: linguistic and nonlinguistic. Linguistic forms of prosody are further divided into those that are morphemic and those that are not. This categorization of prosodic features results in three categories: nonlinguistic prosody, which expresses emotions; linguistic prosody that is nonmorphemic; and linguistic prosody that is morphemic, which is what I define as intonation. This categorization of prosody is unconventional and somewhat controversial, but it facilitates the goal of this book, which is to describe and present evidence in favor of the hypothesis that all meaningful prosody is morphemic.

Hirst (1983: 93) said that “[i]ntonation, what Bolinger has called the ‘greasy part of language’, is notoriously difficult to describe.” ‘t Hart et al. (1990: 2) also recognized the slippery nature of intonation, saying that “it is a fairly elusive subject matter [because it has] features [that] are more difficult to observe, transcribe and analyse than are their segmental counterparts.” Intonation is difficult to study for at least the following four reasons: (1) the term *intonation* may refer to more (or fewer) suprasegmental features and functions when used by different linguists (Johns-Lewis 1985); (2) it is not yet and perhaps never will be possible to mechanically record intonation the way that the native speaker’s ear hears it. Something that a machine records as a rise in pitch, for example, is not necessarily heard by listeners as a rise, and therefore—even though clearly seen on paper—may *not* be linguistically meaningful (Roach 2009); (3) there is no one-to-one correspondence between form and function (‘t Hart et al. 1990; Botinis et al. 2001; Chun 2002); and (4) the various subtypes of suprasegmentals are used simultaneously in speech, one atop another, making it difficult to isolate one type and its associated forms and functions from those of another.

Due to its complex nature, it is not surprising that different linguists have analyzed and described intonation’s forms and functions differently. This book hypothesizes yet another way of analyzing intonation, plus a recategorization of its forms and functions based on that analysis. My proposal is based on the theoretical assumption that intonation is morphemic, and therefore is only justified to the extent that empirical evidence

can be found to support this claim. Before reviewing some of this evidence, however, it will be helpful to first explain in sufficient detail what the hypothesis is. I will do this by first describing what I consider to be the functions of all suprasegmentals, dividing them into the three categories stated above. I will then propose a definition of intonation based primarily on its functions. After that I will then discuss the forms of the suprasegmentals that are used within each of the three categories of prosody.

2.1 The Functions of Suprasegmentals

Botinis et al. (2001: 267) said that “[t]he main functions of intonation are centered round the notions of *prominence*, *grouping* and *discourse*, which are related to various grammatical components as well as linguistic levels,” (emphasis in italics theirs). Referring to intonation, Gussenhoven (2004: 50) said, “people use it to express their feelings; it encodes the information structure of the sentence; [and] it appears sensitive to syntactic categories like ‘argument’ and ‘predicate.’” Many authors, regardless of which linguistic theory they adhere to, seem to agree that intonation is a central part of the grammar, working to mark phrasal, clausal, or theme-rheme boundaries, as well as speech act types, such as question versus statement (Trager 1972; Pierrehumbert and Hirschberg 1990; Crystal 1997a, b; Halliday and Greaves 2008). Except for the marking of speech acts, I will explain below why I do not adopt all of these assumptions about intonation.

Chun (2002) divided the functions of intonation into three categories: (1) grammatical functions, (2) discourse functions, and (3) attitudinal and affective functions. She pointed out, however, that “there are no firmly established or universally agreed upon principles for classifying the functions of intonation” (p. 56), which means that any choices made for delimiting and describing its functions will be somewhat controversial. Nevertheless, it is essential for the purposes of this book that I settle on a particular definition of intonation, which in turn requires that I describe and classify its forms and functions. The categories of intonational functions that are assumed here include only the first two of the three that Chun included—the third one, which is the expression of attitudinal and affective meanings (i.e., human emotions), is assumed to be nonlinguistic and therefore *not* expressed by intonation.

Crystal (1997a, b) recognized only two key functions of intonation by roughly combining Chun’s (2002) second and third functions into a single function. Crystal (1997a: 173) said that in addition to signaling grammatical structure, intonation functions “to express a wide range of attitudinal meanings—excitement, boredom, surprise, friendliness, reserve, and many hundreds more,” and elsewhere he added to this list some “personal attitude[s]: sarcasm, puzzlement, [and] anger” (Crystal 1997b: 202). What Crystal referred to as “attitudinal meanings” included both affective meanings and discourse meanings. In contrast, I use the terms *attitudinal* and *affective* meanings to refer only to nonlinguistic human emotions and distinguish them from the linguistic meanings expressed by intonation, which are assumed to

be part of the lexicon, expressing things such as focus, speaker stance, epistemic and evidential modality, and other discourse-related notions.

An important factor in determining and describing the functions of intonation is to decide how the term *intonation* is used in relation to the term *prosody*; are they the same thing, two separate systems, or is intonation a subset of prosody? The answer to these questions will determine whether they have the same, overlapping, or separate functions. Crystal (1997b), for example, separated intonation from prosody based on form. He said intonation is “the distinctive use of patterns of pitch” (p. 202), while prosody is “variations in pitch, loudness, tempo and rhythm” (p. 313). At the same time, however, he did not seem to separate the functions of these two systems; the meanings he attributed to intonation included the human emotions excitement, boredom, and anger, even though these are expressed by the forms he classified as prosody. Crystal (1997b: 202) listed friendliness, surprise, and anger together under a single function of intonation (i.e., “the communication of personal attitude”) and said that these “can all be signaled by contrasts in pitch, along with other prosodic and paralinguistic features.” This seems to imply that the suprasegmentals that are used to express the discourse notion entailed in surprise and those used to express an emotion such as anger are both expressed by a combination of intonation and prosody. In the present book, in contrast, prosody is defined in a way that clearly separates those forms and meanings that are linguistic from those that are paralinguistic, and intonation is contained within the subset of prosody that is linguist.

This “clear separation” of prosody into two types is largely theoretical; describing and demarcating linguist versus paralinguistic forms and functions of prosody is not made easy simply because we recognize that they should be kept separate. In reference to these two types of suprasegmentals, Ladd (2008: 6) said, “it is a matter of considerable controversy which aspects are which, or whether such a distinction is even possible.” Nevertheless, even if we can only conclude for now that this distinction is *theoretically* possible, proposing theories about the precise nature of prosody and intonation is useful and important for determining how future research might proceed. After Fox (2000: 269) pointed out the frequent attention that scholars have drawn to “the difficulties and uncertainties surrounding [intonation’s] analysis, its systematic description, and its incorporation into linguistic models and theories,” he concluded that “the problems of intonation are more of a theoretical than a practical kind, and relate to its nature and role rather than to its phonetic properties.”

Based on the classifications I adopt here, Crystal’s (1997a, b) list of examples for attitudinal and affective meanings can be divided into three groups: those that express only attitudinal and affective meanings (e.g., boredom, excitement, and anger); those that are a combination of discourse and attitudinal/affective meanings (e.g., surprise, puzzlement, and sarcasm), and all the remaining that are probably too broad to classify (e.g., friendliness and reserve). The first group now lies outside this book’s definition of intonationally-expressed meanings, and instead belongs to nonlinguistic meanings that are expressed through the use of paralinguistic prosody. The labels given to the meanings in the second and third groups should ideally be

replaced by simplified labels that can be clearly identified either as discourse meanings or as attitudinal/affective meanings.¹

Hirst (1977) pointed out that few if any linguists doubt that intonational features contribute information to sentences. The only questions are what kind of information, and whether or not this information is systematic. Is intonational information comprised of discrete features that are acquired by learners along with the other syntactic, semantic, and phonological features of language? If so, we would expect there to be surface differences in the intonational systems of languages. There is plenty of evidence that this is in fact the case. Hirst (1977: 3) went on to say that if, on the contrary, “we consider intonation as merely a direct, physical manifestation of the speaker’s emotions and feelings, [then] we should normally expect different languages to use the resources of intonation in very similar, if not exactly the same, ways.” Suprasegmentals can be used in both of these ways, i.e., to express discrete linguistic features *and* to express emotions. The task at hand then is to separate and classify these two types of suprasegmentals. Some linguists include the expression of emotions as a property of intonation (e.g., Chun 2002; Gussenhoven 2004), but here these are instead classified as a property of nonlinguistic prosody.

Many authors have distinguished suprasegmentals that are linguistic from those that are nonlinguistic. The latter have been referred to as a form of animal communication that functions to express emotions and nonlinguistic pragmatic meanings. Couper-Kuhlen (1986: 174) said, “we must distinguish an unmonitored, purely physiologically determined externalization of emotional state, presumably universal across linguistic communities, from a ‘cognitively’ monitored expression of attitude, conventionalized and communicative in purpose.” Fox (2000: 270) likewise distinguished “non-linguistic” suprasegmentals that relate to emotions and attitudes from “the pitch features associated with [linguistic functions and intonation patterns that are] by no means always ‘natural’ and universal, but differ from language to language, and hence reflect an arbitrariness characteristic of linguistic, rather than non-linguistic, phenomena.”²

Gussenhoven (2004: 50) distinguished between two categories of intonation, saying that “intonation is both a form of animal communication . . . and part of the linguistic structure.” He said that human language has the arbitrariness of the sign and that some aspects of suprasegmentals are clearly nonarbitrary because, across languages,

[w]hen we are excited, our pitch goes up, and when we are depressed we tend to have a low pitch with few excursions . . . When we wish to emphasize a word, we may raise our pitch, in addition to raising our voice in the sense of speaking more loudly. When we want to sig-

¹Attentive readers may note that I later use the terms *surprise* and *doubt* to describe some of the discourse meanings within my own research in Chap. 6. This does not conflict with what I say here about such terms, however, because I use those terms only in reference to clearer, fuller definitions; I do not use them to refer directly to the intonational forms themselves.

²Note that Couper-Kuhlen (1986) and Fox (2000) both use the term “attitude,” but the former uses it to describe a linguistic meaning, while the latter uses it to describe nonlinguistic meaning, illustrating yet another example of inconsistency in the use of terminology.

nal—for real, or more probably in jest—that we need the speaker’s protection or deserve his mercy, we instinctively raise our pitch, to create a “small” voice. (Gussenhoven 2004: 51)

We can see from this that Gussenhoven (2004), like Crystal (1997b), distinguished between two forms of suprasegmentals. The forms that Crystal referred to as prosody were considered by Gussenhoven to be a form of animal communication. And the forms Crystal referred to as intonation, Gussenhoven considered as part of the linguistic structure. However, Gussenhoven placed both of these types of suprasegmentals under the category “intonation,” as did Pike (1945: 24), who said that “various types of intonation, such as the general pitch of the voice as a whole in contrast to the different pitches occurring within a single sentence, must be studied separately in so far as is possible.” What I call paralinguistic prosody is what Pike referred to as pitch of the voice as a whole (i.e., its range and key), and what I call linguistic prosody is what Pike called sentence-internal pitch manipulation. Other authors have also followed Gussenhoven’s (2004) and Pike’s (1945) classification of referring to intonation as both linguistic and paralinguistic (e.g., Couper-Kuhlen 1986; Brazil 1997; Fox 2000). In contrast, I divide suprasegmental features of speech (i.e., prosody) into two categories: (1) linguistic suprasegmentals and (2) paralinguistic suprasegmentals. Category 1 is assumed to be part of the linguistic structure, and category 2 a form of animal communication. Category 1 is further divided into morphemic and nonmorphemic suprasegmentals, with the former being what I define as intonation (see Table 2.1). This definition of intonation is narrower than that of Pike (1945) and Gussenhoven (2004) because it excludes those forms that are not part of the linguistic structure. In fact, it is narrower than most linguists’ definition of intonation because it further excludes those forms of prosody that have been said to delimit phrasal structures—my reasons for this exclusion are given below.

Some linguists make a clear distinction between intonation and prosody, others partially combine the two, and yet others use the two terms interchangeably. Ladd’s (2014, Chap. 3) detailed description of the origin and historical uses of the term

Table 2.1 Suprasegmental categories based on function

Prosody (all suprasegmental features of speech)		
Linguistic suprasegmentals		Paralinguistic suprasegmentals
Intonation	Nonmorphemic linguistic prosody	Non-linguistic prosody (a form of animal communication)
Functions: 1. To express grammatical functions 2. To express discourse meanings	Functions: • To produce lexical tone • To mark the prosodic structure	Function: • To express attitudinal, affective meanings

prosody indicates that most authors consider intonation to be a subcomponent of prosody. Johns-Lewis (1985: xix) asked, “Is there a dividing line between intonation and prosody? The answer, as with so many terms, is that it depends on who is using the terms.” Like most authors, she classified intonation as a subset of prosody, and I adopt that practice here, additionally restricting it to being a subset of linguistic prosody. This results in three subsets of prosody with some overlapping forms (see Sect. 2.3), but theoretically these different types of prosody are qualitatively different in nature because they have no overlapping functions, as indicated in Table 2.1.

Some of the classifications in Table 2.1 are nonconventional, so they warrant clarification. One classification that should not be controversial is making intonation a subset of prosody. Linguists must choose whether to classify these as the same thing, to classify them as two separate things, or to classify intonation as a subset of prosody. I have chosen the latter for two reasons: (1) prosody has multiple functions, so it makes practical sense to distinguish it from intonation, which can then be used as a term to refer to a subgroup of prosody’s functions and (2) classifying intonation as a subset of prosody, rather than as something distinct, makes sense because intonation is made up of prosodic features and therefore should be considered as belonging to prosody. Inherent features are associated with vowels and consonants and can be defined without reference to the sequence of sounds in an utterance. Prosodic features, in contrast, can only be defined in reference to acoustic changes or contrasts within the utterance, or in reference to a person’s voice range (Ladd 2008: 189–92). What is unusual about Table 2.1 in this respect is that I classify all prosodic features of language as prosody, including tonal morphemes and lexical tones, which, as far as I know, have never before been referred to as forms of prosody and/or intonation.

Gussenhoven and Jacobs (2014: 148–9) classified tones into three types based on their functions: *lexical tones*, which distinguish syllabic morphemes which otherwise share the same segments; *grammatical tones*, which are themselves morphemes (i.e., what I refer to as tonal morphemes); and *intonation tones*, which “function to signal discourse meaning or phrasing.” They said that tonal morphemes, “unlike intonational morphemes, have meanings that fit into the morphological and syntactic paradigms of the language, instead of expressing discursual meanings” (p. 157). This differs from what I am arguing here, which is that discourse meanings are in the lexicon and that intonational tones therefore belong in the morphological and syntactic paradigm of language to the same extent as tonal morphemes. This means, for example, that a rising question tone in English is analyzed no differently from the rising tone that marks progressive aspect on verbs in Inland Ewe, a Kwa language from West Africa (Aboh and Essegbey 2010). These two tones belong to different categories, with one being a discourse particle and the other an aspect particle, but both are equally represented in the lexicon and in the syntax. As such, they are both analyzed here as tonal morphemes.

The first function listed for intonation in Table 2.1 is also controversial. The expression of grammatical functions refers to what are traditionally called tonal morphemes, like the example from Inland Ewe just mentioned. (It must be made clear that the expression of grammatical functions does *not* refer to the use of intonation to mark phrasal structure, which I tentatively argue is not a function of into-

nation.) The production of lexical tone is shown as a function of nonmorphemic linguistic prosody because lexical tone is linguistic, has prosodic features, and does not comprise morphemes.³ The other function listed under nonmorphemic linguistic prosody is the marking of prosodic structure. Many linguists consider the marking of phrasal, clausal, and prosodic boundaries to be a function of intonation. I instead adopt the view of Ladd (2008), who said,

intonation has no privileged status in signaling prosodic structure ... intonational features of pitch and relative prominence are distributed in utterances in ways allowed by the prosodic structure. In some cases this means that conspicuous phonetic breaks occur at major constituent boundaries, but this is neither the essence of the boundary nor the only factor governing the distribution of the intonational features. (Ladd 2008: 10–11)

It must be pointed out that prosodic structure does not fully match syntactic structure; it is merely related to it. In her study on the interactions between syntax and prosody in Connemara Irish, Elfner (2012) argued that the prosodic form of a sentence is not only a result of syntactic structure; it is also influenced by linearization and prosodic well-formedness. Bennet and Elfner (2019: 153) took it as a “fact that prosodic structure is derived from syntax, but need not be identical to it.” Explaining why this is the case, Gussenhoven and Jacobs (2014: 203) said that “[s]ince morphosyntactic constituents of a given rank may vary hugely in length, a one-to-one correspondence between phonological and morphosyntactic constituents would put unreasonable demands on speakers.” It makes sense that prosodic structure should not clash with morphosyntactic structure, as that would affect communication, but the fact that they are not the same makes it reasonable to assume that the morphosyntactic structure is not determined (i.e., marked) by the prosodic structure but is instead marked by lexical and grammatical features. I tentatively assume that prosodic structure is strictly phonetic in nature, residing outside the lexicon and syntax; it is merely a by-product of the grammatical structure of language (along with linearization and prosodic well-formedness). Based on this assumption that the suprasegmentals used for prosodic phrasing are nonmorphemic, the marking of prosodic structure is not listed as a function of intonation in Table 2.1. Saying that the marking of prosodic structure is not a function of intonation is no trivial matter, and it has consequences for the theory of intonation that I am proposing here. The implications and some of the issues involved are discussed in Sect. 7.3.

Suprasegmentals that mark prosodic structures are distinguished from those that have discourse-related meanings. In the case of topicalization, for example, an associated tone can be seen as marking the topicalized phrase in the same way that topic particles do in languages like Chinese and Japanese. Consider this observation from Gussenhoven and Jacobs (2014):

³ Categorizing lexical tone under the term linguistic prosody makes sense for these reasons, but it does not mean I endorse the practice of regularly referring to lexical tone as prosody. That would be impractical; thinking of it and referring to it as distinct from prosody facilitates discussing how lexical tones interact with what is traditionally thought of as prosody.

A phrasing function of tone occurs in an English sentence like *Once we're in China, we can practice our Chinese*, where the last syllable of *China* is likely to have a high tone indicating the boundary. When tones function to signal discursal meaning or phrasing they are intonational tones. (Gussenhoven and Jacobs 2014: 149)

This tone on the second syllable of “China” can reasonably be analysed as a tonal morpheme functioning as a topic marker. Jumping ahead to the ideas that will be discussed in Chap. 7, the prepositional phrase “once we’re in China” is a topicalized phrase that is assumed to have moved to what is called the specifier position of a Topic phrase at the left periphery of the sentence. This Topic phrase is headed by a feature [+Top] (cf. Aboh 2010). The tone on the second syllable of “China” can be analyzed as the phonological realization of this [+Top] feature, making it a tonal morpheme that is no different from the segmental topic markers in other languages (e.g., Japanese *wa*; Korean *nun*; Mandarin *ne*). The only difference is that instead of being segmental, it is a floating tone that is associated with one or more of the segments at the end of the topicalized phrase. Sometimes no tone is used, and the presence of the [+Top] feature is made known only by the fact that the topicalized phrase has raised. This is not surprising since topic markers in Chinese are also optional. Another example comes from Aboh (2016), who contrasted a Hungarian sentence with one from Gungbe, both of which included both a focused and a topicalized phrase. He said the only difference between their ways of marking these phrases is that Hungarian uses intonation where Gungbe uses the segmental topic marker *yà* and a focus marker *wě*. This shows a clear contrastive comparison between segmental and intonational topic and focus markers, indicating that the intonational forms associated with topicalized and focused elements are morphemic. This differs from those tones associated with prosodic structure that do not appear to have any segmental counterparts.

2.2 A Definition of Intonation Based on Its Functions

Based on the classifications shown in Table 2.1, a definition of intonation is proposed as follows:

(1) *Intonation*:

A suprasegmental form that has semantic content or a grammatical function

The definition in (1) incorporates the classic definition of a morpheme: “A morpheme—the minimal linguistic unit—is thus an arbitrary union of a sound and a meaning (or grammatical function) that cannot be further analyzed” (Fromkin et al. 2013: 38). Defining intonation as morphemic results in the normal practice of excluding lexical tone, which is classified here as a form of nonmorphemic linguistic prosody. However, this definition of intonation now includes tonal morphemes, which have not traditionally been considered forms of intonation. They have instead been referred to as suprasegmental morphemes that function to mark things such as grammatical aspect, definiteness, grammatical case, and so on. These have always been seen uncontroversially as morphemes that reside in speakers’ lexicons and that

are therefore part of core syntax and semantics, but they have, as far as I know, always been described and defined without reference to intonation beyond how they may interact with intonational phonologically. The definition in (1) also defines the intonation used to express discourse meanings as tonal morphemes. The idea that discourse tones are morphemic in the same sense as grammatical tonal particles will be discussed in Chap. 3, and further evidence will be presented in Chaps. 4 and 6.

The definition in (1) will now be compared with some definitions found in the literature in order to illustrate the differences. Starting with a relatively short and simple definition, Trench (1996: 2) said that “intonation is the linguistic use of pitch in utterances.” There are two key differences between his definition and that of (1): first, Trench too narrowly defines intonation’s form as pitch alone (see Sect. 2.3) and second, in Trench’s own words, it “specifies that intonation is concerned with utterances” (p. 3). While specifying that intonation relates to utterances works to exclude lexical tone, as it should, it is not specific enough—it does not work to clarify what functions and meanings are included or excluded.

Cruttenden (1997: 7) said that “intonation involves the occurrence of recurring pitch patterns, each of which is used with a set of relatively consistent meanings, either on single words or on groups of words of varying length.” Again this definition takes the practical step of simplifying the form of intonation to pitch alone, but a more critical difference between this definition and the definition in (1) is that intonational meanings are said by Cruttenden to be only “relatively consistent.” This does not align with the strong claim entailed in (1), which is that intonational forms have semantic content or a grammatical function—their core meanings are therefore assumed to be constant from one occurrence to the next.

Gussenhoven (2004: 12) said “intonation is treated as the use of phonological tone for non-lexical purposes, or—to put it positively—for the expression of phrasal structure and discourse meaning.” The term *phonological tone* implies the possibility of a combination of phonetic features working in tandem to form a tone, rather than pitch alone. The definition in (1) therefore could have used the term *phonological tone* in place of “suprasegmental form.” Saying that intonation is used for nonlexical purposes works to exclude lexical tone, as it should, but it also excludes tonal morphemes, thus excluding everything that I define as intonation. Another difference is that Gussenhoven’s definition includes the expression of phrasal structure as a function of intonation, while I tentatively exclude this for the reasons explained above.

Ladd (2008: 6) defined intonation as “the use of suprasegmental phonetic features to convey ‘postlexical’ or sentence-level pragmatic meanings in a linguistically structured way.” Saying that intonation is “postlexical” excludes lexical tone, but it again also excludes tonal morphemes, and saying that intonation expresses “sentence-level pragmatic meanings” excludes all word-level or phrase-level grammatical tones. While intonation has traditionally only included sentence-level tones, there is no reason that such tones cannot be categorized together with word-level or phrase-level grammatical tones. Discourse tones have frequently been compared with segmental discourse particles, so categorizing all tonal morphemes together is analogous to categorizing all segmental particles together. The differences between a tonal discourse particle and a tonal grammatical particle are their meaning, their

scope, and their grammatical function. Categorizing them together under the term *intonation* is comparable with categorizing all of their segmental counterparts, both grammatical and discursal, under the term *particles* for the purpose of describing and categorizing linguistic features. In both cases, we recognize that the collection of grammatical elements included under an umbrella term can and should be sub-categorized. Another difference between the definition in (1) and Ladd's (2008) definition is that the former considers intonational forms to have meaning beyond merely pragmatic meaning (see Chap. 3).

2.3 The Forms of Suprasegmentals

The paralinguistic expression of emotional attitudes is a form of animal communication which uses prosodic forms that are basically universal across languages, and these affective meanings are not lexicalized in any language. Calling them a form of animal communication is not meant to imply that the nonlinguistic meanings they express are not unique to humans. The paralinguistic suprasegmentals that accompany our speech are used to express uniquely human emotions, and their forms are a unique product of the human vocal tract. They are not part of the syntax or lexicon of human language, so it is reasonable to conclude that they are a form of animal communication that is used by the animal species known as *Homo sapiens*. Their forms and meanings are naturally rather consistent across the species, and many of their qualities may or may not be shared with the forms of communication used by other animal species.

Saying that the forms of paralinguistic prosody are basically universal does not mean one would expect it to be used in the same way cross-linguistically and cross-culturally. From language to language, there will be some variations in both the forms and the meanings of paralinguistic prosody. Variations of form will occur because prosodic forms overlap and interact with the unique phonological features of a given language, including its systems of intonation and linguistic prosody. And differences of both form and meaning will result from cultural and speaker individual differences that influence the expression and interpretation of human emotions.

Some evidence for the universality of "paralinguistic features" comes from Maekawa (2004: 8), who concluded that "the perception of [paralinguistic information] as voice-quality is language-independent, or universal, like perception of emotion,⁴ while the perception of [paralinguistic information] as manifested by the manipulation of the features of phrase-phonology is language-dependent." He came to this conclusion based on his experiment with various forms of suprasegmental information, all of which he called "paralinguistic information," but which I refer to as prosody. His term "voice quality" referred to what I call paralinguistic prosody.

⁴It should be noted that the universality of the perception of emotion through facial expressions has been challenged, e.g., Russell (1994).