The Structure of Spoken Language

Using an innovative approach, this book focuses on a widely debated area of phonetics and phonology: intonation, and specifically its relation to metrics, its interface with syntax, and whether it can be attributed more to phonetics or phonology, or equally to both. Drawing on data from six Romance languages (French, Italian, Spanish, Portuguese, Catalan, and Romanian), whose rich intonation patterns have long been of interest to linguists, Philippe Martin challenges the assumptions of traditional phonological approaches, and re-evaluates the data in favor of a new usage-based model of intonation. He proposes a unified description of the sentence prosodic structure, focusing on the dynamic and cognitive aspects of both production and perception of intonation in speech, leading to a unified grammar of Romance languages sentence intonation. This book will be welcomed by researchers and advanced students in phonetics and phonology.

PHILIPPE MARTIN is a Professor in the Linguistics Department at the Université Paris Diderot.

The Structure of Spoken Language

Intonation in Romance

Philippe Martin

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781107036185

First published 2015

Printed in the United Kingdom by TJ International Ltd. Padstow Cornwall

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data
Martin, Philippe, 1944– author.
The structure of spoken language : intonation in Romance / Philippe Martin.
pages cm
ISBN 978-1-107-03618-5 (hardback)
1. Romance languages – Phonetics – Intonation. 2. Romance languages – Phonology. 3. Romance languages – Phonology, Historical. 4. Romance languages – Spoken Romance languages. 5. Intonation (Phonetics) 6. Biolinguistics. I. Title.
PC81.5.M27 2015
440′.0415–dc23

2015012063

ISBN 978-1-107-03618-5 Hardback

Contents

List of figures and maps

Figures

List of tables

Preface

This book is the culmination of some forty-five years of personal research on intonation. When I first went to Canada in 1968, I had the rare and precious opportunity to be hired by Pierre Léon (1926–2013), then a young and enthusiastic phonetician, who was eager to carry out research on all possible aspects of intonation, in didactics, phonostylistics, and phonology. Being a young graduate in electronic engineering, I was given as a first task the job of developing an acoustic instrument capable of measuring in real time the fundamental frequency of speech (an acoustic measure of vocal folds vibration frequency). Being quite new to the field, I thought this assignment would be easy to complete in some two or three weeks! I realize today, after all these years, that despite the considerable progress that has been made, the question of fundamental frequency tracking has still not been completely solved. In fact, new generations of young specialists in speech signal analysis are regularly tackling the problem, only to discover that even though systems are becoming more reliable, there is always room for improvement. The grail in the field would be the availability of an algorithm which would give reliable results for all cases of speech recording conditions, or at least all cases where a human listener can perceive the melodic variations resulting from fundamental frequency change.

Nevertheless, I had no fear at that time, especially as I had received a brand new PDP 8/I for my exclusive usage (the PDP 8/I was at that time a revolutionary laboratory computer with 8 kBytes central memory and a 32 kBytes hard disk). I spent all my work and leisure time developing a hybrid analog-digital system, capable of delivering in real time a fundamental frequency curve with an extended range (70 Hz to 500 Hz). The curve was displayed on a high remanence oscilloscope screen and filmed on a large TV screen. At a time when all intonation acoustical data were obtained either with a spectrograph (limited to 2.4 second analysis and requiring 3 to 5 minutes of processing, not to mention the necessary and tedious manual setup) or with a mingograph (hooked to a not too reliable voice filter operating in a limited frequency range and requiring an expensive roll of ink or UV photographic paper to print the results), this was an important achievement which had direct consequences for my own prosodic research.

Indeed, I was then able to process a very large amount of data, essentially in French, sometimes playing with the intonation of my own voice and getting the resulting melodic curves immediately, and was curious to see if some pattern would emerge from all these trials. In the late months of 1973, I finally got an idea pertaining to the frequently observed regularity of F0 patterns, an idea that I formalized with the term *contraste de pente* (contrast of melodic slope) in French, and published in the review *Linguistics* in 1975 (Martin, 1973, 1975). At the same time, I coined the terms *structure prosodique* and *hiérarchie prosodique* (prosodic hierarchy and prosodic structure) to name a hierarchical organization of minimal prosodic units or prosodic words, containing one and only one prosodic event indicating this hierarchy. Although referring to the same kind of experimental data in French, papers published at that time (e.g. Vaissière, 1975; Émerard, 1977) were phonetically rather than phonologically oriented but brought comparable data.

To my regret, my *Linguistics* paper had almost no impact on the research domain in prosody, except in France. Only five years later, however, papers using the term prosodic structure appeared, but unfortunately without ever mentioning my earlier work. Later, speech analysis computer software became popular (Signalyze, WinPitch, Praat, and so on), and phonologists (essentially based in US universities) were happy to discover a new playground. To differentiate their activity from that of the phoneticians, who were not considered seriously by linguists at the time, and to avoid being confused with them, they called it "laboratory phonology," which corresponds to what most phoneticians have actually been doing for a century or more.

The purpose of this monograph is to present an alternative theoretical approach that attempts to describe and understand the prosodic organization of sentences. In this endeavor, I briefly present a critical exposition of the main aspects of the dominant Autosegmental-Metrical model (henceforth AM), succinctly describing existing research using this approach for Romance languages such as European French, Italian, Spanish, Catalan, European Portuguese, and Romanian (Post, 2000; D'Imperio 2002; D'Imperio et al., 2005; Michelas & D'Imperio, 2010; Sosa, 1999; Hualde, 2003; Prieto, 2014; Frota, 2009). Then I introduce an alternative model, called Incremental Storage Concatenation (henceforth ISC) derived from the Storage-Concatenation model I proposed in 2009 (Martin, 2009). In this model, I highlight some characteristics, apparently never mentioned in AM descriptions, formalized as a set of constraints limiting the number of prosodic structures that could be associated with a given text.

This leads to a concept of intonation that from the start completely dissociates sentence text from its hierarchical organization by syntax. This concept departs dramatically from earlier concepts of prosodic structure

conceived under the AM approach, where only one such structure can be associated with a given syntactic structure, even if its restructuration appears possible,in order to obtain a better eurhythmicity (Post, 1999).

The set of constraints, originally part of the Storage-Concatenation framework, i.e. planarity, the seven syllables rule, eurhythmicity, stress clash, and syntactic clash, made me look for an underlying explanation that gives a proper account for the observed constraints. A key aspect is their time dimension and especially the dynamic process performed by listeners to recover the prosodic structure intended by the speaker or the writer. Examining the consequences of the time domain aspect of the process is the key to a better understanding of the observed data, an aspect that is often neglected or totally ignored in the current literature. Indeed, the usual reasoning on a two-dimensional plane of a sheet of paper limits considerably an understanding of the mechanisms necessarily used by the listener in the perception of the prosodic structure.

Pushing this exploration further, I related this model to results obtained recently in the neurolinguistic domain, and particularly those concerning evoked potential linked to prosodic stimuli. These results lead me to propose a new and coherent model based not only on the time dynamics of the prosodic structure but also, and perhaps even more interestingly, on specific cognitive mechanisms, in particular those involving short-term memory (Gilbert, 2012). This approach suggests a convincing set of explanations pertaining not only to the set of constraints relative to the prosodic structure but also to some phonetic data, such as the duration of minimal units of prosody (defined below as prosodic words), the minimal and maximal time interval between consecutive stressed syllables, and even the speed limits of silent reading.

The second part of this book is devoted to applications of the model presented in the first part to the analysis of data in some Romance languages, starting with French, often considered as the ugly duckling among other languages of the same family as it is deprived of lexical stress. This second part itself is divided according to the type of data analyzed: read/laboratory speech and spontaneous/non-prepared speech. In this latter set of chapters, I use a modified macrosyntax approach derived from the GARS (Groupe Aixois de Recherche en Syntaxe) work (Blanche-Benveniste, 1990, 2000) for both the text and the prosodic aspects of speaker productions.

I sincerely hope that this book will help both new and experienced researchers in the field of prosody to restore sentence intonation to its deserved place in linguistic studies. I will try to show that far from being the cherry on the phonological cake for some, intonation is the essential linguistic base for both speech production and speech perception.

Acknowledgments

I have many people to thank, and in the first place, Pierre Léon (1926–2013) who, like Obelix, a cartoon character in the adventures of the French comic book *Astérix*, "plunged me in a barrel of prosody when I was little." From Pierre Léon I learned a lot of facts about intonation in linguistics, stylistics, phonetics, etc., and about how to survive in the academic world.

In addition, I had the privilege to meet and work with the outstanding linguist Claire Blanche-Benveniste (1935–2010). She had a tremendous influence on my research, always encouraging me to improve in our countless fruitful and pleasant discussions.

Many other people helped me in various ways. In particular, I would like to thank (in alphabetic order):

Mathieu Avanzi (Université de Neuchatel) for his numerous useful (and exacting) comments;

Helen Barton (Cambridge University Press) for her constant support and encouragement in this project;

Gabriela Bilbiie (Université Paris Diderot) for her help in elaborating the Romanian corpus;

Victor Boucher (Université de Montréal) for his original and fruitful views on speech perception and his constant support for this project;

Georges Boulakia (Université Paris Diderot) for his constant friendship and understanding;

Marie Claude Capt-Artaud (Université de Genève), formerly skeptical but now convinced;

Emanuela Cresti (Università degli Studi di Firenze) for our discussions and her constant friendship;

Jeanne-Marie Debaisieux (Université Paris 3) for the trust she placed in my research;

Élisabeth Delais-Roussarie (Université Paris Diderot) for many interesting discussions;

Didier Demolin (Université de Grenoble) for his indefectible friendship;

José Henri Deulofeu (Université Aix-Marseille) for teaching me what macrosyntax is and staying my friend;

Helena Dowson (Cambridge University Press) for her patience and encouragement;
Caterina Falbo (Università degli Studi di Trieste) for her understanding and inspiration;
Ana Maria Fernández Plana (Universitat de Barcelona) for her help in elaborating and recording the Catalan corpus;
Jacqueline French (Cambridge University Press) for her tireless efforts to improve the text;
Aline Germain (Middlebury College) for her trust in my theories;
Annie Gilbert (Université de Montréal) for her illuminating view on syllables;
Rémi Godement (Université Paris Diderot) for his trust in being one of my doctoral students;
Sarah Green (Cambridge University Press) for her constant help with this project;
Jane Leung (University of Toronto) for her patience and impatience (at times), and her moral support;
Mirian Matta-Machado (Universidade Federal Fluminense) for her constant friendship and support;
Massimo Moneglia (Università degli Studi di Firenze) for his friendship and exceptional cuisine;
Patricia Perez (Université Paris Diderot) for her friendly participation in corpora building;
Michaela Pirvulescu (University of Toronto) for her help in elaborating the Romanian corpus;
Guan Qianwen (Université Paris Diderot) for interesting discussions about Mandarin prosody;
Tommaso Raso for his constant support and friendship;
Mario Rossi (Université de Provence) for his trust in my software realizations;
Alexandre Sévigny (University McMaster) for his support in difficult days, and brilliant suggestions in incremental syntax;

And finally, I would like to thank all my present and past students (master's and doctoral), who, by their presence and active participation during my lectures, encouraged me to write this book, and, of course, all the friends and colleagues I may have inadvertently forgotten.

Every effort has been made to secure necessary permissions to reproduce copyright material in this work, though in some cases it has proved impossible to trace copyright holders. If any omissions are brought to our notice, we will be happy to include appropriate acknowledgments in any subsequent edition.

Key concepts

To help the reader to quickly evaluate the distance from known (and dominant?) concepts in the field of intonation studies, the following list contains the essential nonstandard theoretical points developed in this book.

1. This book is about the structure of spoken language.
2. Spoken language is made of time sequences of syllables organized into stress groups (basic units of speech are syllables, not phonemes).
3. Stress groups are not necessarily aligned with words or syntactic groups; however, they are aligned on complete words (i.e. their beginnings and ends are aligned on beginnings and ends of lexical units – words).
4. Stress groups are also called rhythmic groups, accent groups, prosodic words, and Accent Phrases in the literature.
5. Prosodic words are segments of prosody associated with and aligned on stress groups.
6. Prosodic words are organized hierarchically by a prosodic structure.
7. Specific prosodic markers indicate prosodic structures; they allow the listener to reconstitute dynamically the speaker intended prosodic structure in an incremental time fashion.
8. Prosodic markers are instantiated by prosodic events located on stressed groups' stressed and final vowels.
9. Prosodic events are instantiated by prosodic contours, described primarily in acoustic terms of duration, melodic contour, and intensity.
10. (Silent) reading and speaking are described as an Incremental Storage-Concatenation (ISC) process.
11. Recovering the prosodic structure in (silent) reading mode is a specific process distinct from listening to speech.
12. Generation of spontaneous speech involves chunks of prosody hosting syntactic constructions, which in turn host morphological units.
13. There is therefore a precedence of prosody over syntax, and of syntax over morphology.
14. It follows that the same prosodic structure can host various syntactic and morphological constructions, i.e. different texts.
15. Conversely, more than one prosodic structure can be associated with a given text (for example when reading).

16. Prosodic boundaries (between an AP, ip, or IP) do not correspond necessarily to syntactic or macrosyntactic boundaries. Likewise, (macro)syntactic boundaries do not necessarily correspond to prosodic boundaries.
17. Stress shift in stress clash conditions entails a reallocation of stress groups organized hierarchically in the prosodic structure;
18. Generation of a prosodic structure when reading involves the precedence of syntax (analyzed by the reader from the written text).
19. Prosodic structures are not necessarily congruent with the sentence syntactic structure. They do not result from restructuration of the prosodic structure either. Actually they do not coexist with syntax; they precede syntax.
20. Prosodic markers are subject to neutralization of some of their acoustic features when partially or totally redundant in a given prosodic structure configuration.
21. Prosodic markers must be acoustically similar in their respective domain.
22. Acoustic features describing prosodic events ensure a necessary and sufficient differentiation between prosodic markers (melodic contours) in the prosodic structure.
23. Prosodic structures are constrained by a set of rules: planarity / seven syllables / stress clash / syntactic clash / eurhythmy (the latter for read speech).
24. Neurocognitive properties and processes may explain these constraints.
25. Prosodic structure and prosodic markers properties are extended to macrosyntax.
26. Broad and narrow focus are subcases of macrosyntax configurations (Prenucleus, Nucleus, Postnucleus).
27. There is a macrosyntax analysis of sentence intonation (no Prefix, only prosodic Nucleus, prosodic Parenthesis, and prosodic postfix).

In order to be compatible with the many other studies on intonation that are probably familiar to most readers, I use the terms R (prosodic structure root), IP (intonation phrase), ip (intermediate intonation phrase), and AP (Accent Phrase) throughout this book whenever possible, despite potential general conceptual differences.