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.
Contents

List of figures and maps x
List of tables xix
Preface xxi
Acknowledgments xxv
Key concepts xxvii

1 Introduction 1
   The respiratory cycle 1
   The source-filter model of phonation 3
   Emotions 5
   Voiced and unvoiced speech sounds 7
   Laryngeal frequency 7
   Fundamental frequency and melodic curve 7
   Intensity 9
   Spectrographic analysis 9
   Syllabic duration 10
   Syntax and prosody 11
   The prosodic structure: the structure of spoken language 13
   Stressed syllables 13
   Intonation and syntax 14
   Brain waves and prosody 14
   A Copernican change 15
   From laboratory to spontaneous speech 16
   Reading and listening 16
   Romance languages 17

2 The role of technological advances 20
   The kymograph 20
   The spectrograph 21
   Fundamental frequency tracking 23
   First results 23
   Electroencephalography and brain waves 27
   Transcription and alignment of speech 27

3 Transcription systems 29
   Acoustic and perceived data 30
Contents

Obtaining data: pitch curves 30
Selecting data 32
  Historical background 32
  The AMPER project 36
  The Prosogram 36
  ToBI 38
  INTSINT and Momel 40
  Analor 41
Transcription as theory 42
Perception and interpretation 43
  A phonological transcription system 44

4 The Autosegmental-Metrical Prosodic Structure 46
  A brief description 46
  Properties 48
  Applying the concept 51
  Questions and remarks 54
  The prosodic structure revisited 56

5 The Incremental Prosodic Structure 59
  Melodic curves 59
  The stress group 61
  The prosodic word 62
  Syllabic chunking 63
  The time dimension 64
  Conversion of syllabic chunks 65
  The syllable in the stress group 66
  The stress group in the sentence 68
  Classes of conclusive contours 68
    Basic modalities 68
    Modality variants 69
    Alternative questions 71
  Iconicity of conclusive contours 71
    Imperative contour 73
    Implicative contour 73
    Contour of surprise 74
    Contour of doubt 75
  The Incremental Prosodic Structure 76
    Independence 79
    Prosodic events 79
    Properties 81
    Prosodic phrasing 82
    Planarity 83
    Connexity 83
    Domain 84
    Neutralization 85
  Differentiation in the time domain 85
    Differentiation of prosodic events 86
    The dynamic prosodic structure 87
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Incremental Storage-Concatenation process</td>
<td>88</td>
</tr>
<tr>
<td>Preplanning</td>
<td>90</td>
</tr>
<tr>
<td>Melodic contours features</td>
<td>90</td>
</tr>
<tr>
<td>One prosodic word</td>
<td>91</td>
</tr>
<tr>
<td>Two prosodic words</td>
<td>92</td>
</tr>
<tr>
<td>Three prosodic words ended with C0</td>
<td>92</td>
</tr>
<tr>
<td>Contrast of melodic slope</td>
<td>94</td>
</tr>
<tr>
<td>Three prosodic words ended with C1</td>
<td>94</td>
</tr>
<tr>
<td>Prosodic structure constraints</td>
<td>96</td>
</tr>
<tr>
<td>The <em>arc accentuel</em> in French</td>
<td>96</td>
</tr>
<tr>
<td>Stress clash</td>
<td>97</td>
</tr>
<tr>
<td>Minimum duration of prosodic words</td>
<td>98</td>
</tr>
<tr>
<td>Maximum duration of prosodic words</td>
<td>99</td>
</tr>
<tr>
<td>Eurhythmy</td>
<td>101</td>
</tr>
<tr>
<td>Word alignment</td>
<td>104</td>
</tr>
<tr>
<td>Syntactic clash</td>
<td>105</td>
</tr>
<tr>
<td>Experimental data</td>
<td>106</td>
</tr>
<tr>
<td>Brain waves and prosodic structure</td>
<td>107</td>
</tr>
<tr>
<td>Theta brain waves and the perception of syllables</td>
<td>107</td>
</tr>
<tr>
<td>Delta brain waves and stressed syllables</td>
<td>108</td>
</tr>
<tr>
<td>Delta brain waves frequency range</td>
<td>110</td>
</tr>
<tr>
<td>Prosodic structure constraints and brain waves</td>
<td>111</td>
</tr>
<tr>
<td>Stress groups and brain waves</td>
<td>113</td>
</tr>
<tr>
<td>Constraints revisited</td>
<td>113</td>
</tr>
<tr>
<td>Sequential sentence structuration by prosody and syntax</td>
<td>115</td>
</tr>
<tr>
<td>A simple example: telephone numbers</td>
<td>116</td>
</tr>
<tr>
<td>6 Lexical stress in Romance languages</td>
<td>120</td>
</tr>
<tr>
<td>Stress and accent</td>
<td>120</td>
</tr>
<tr>
<td>Stress in various languages</td>
<td>121</td>
</tr>
<tr>
<td>Stressed syllables in Latin</td>
<td>122</td>
</tr>
<tr>
<td>Stressed syllables in Romance languages (other than French)</td>
<td>123</td>
</tr>
<tr>
<td>Orthographic convention and homographs</td>
<td>124</td>
</tr>
<tr>
<td>Rules for word stress placement</td>
<td>126</td>
</tr>
<tr>
<td>A statistical approach</td>
<td>127</td>
</tr>
<tr>
<td>A phonological-phonetic approach</td>
<td>127</td>
</tr>
<tr>
<td>A phonological approach</td>
<td>128</td>
</tr>
<tr>
<td>A morphophonetic approach</td>
<td>128</td>
</tr>
<tr>
<td>A morphological approach</td>
<td>128</td>
</tr>
<tr>
<td>French</td>
<td>130</td>
</tr>
<tr>
<td>Secondary accent and <em>arc accentuel</em></td>
<td>131</td>
</tr>
<tr>
<td>The <em>groupe de sens</em>...</td>
<td>131</td>
</tr>
<tr>
<td>Stress variations in Romance languages</td>
<td>132</td>
</tr>
<tr>
<td>7 The Incremental Prosodic Structure in six Romance languages</td>
<td>133</td>
</tr>
<tr>
<td>EuRom4 and EuRom5</td>
<td>134</td>
</tr>
<tr>
<td>The process of reading</td>
<td>135</td>
</tr>
<tr>
<td>Note on figures</td>
<td>136</td>
</tr>
<tr>
<td>The Incremental Storage-Concatenation process</td>
<td>136</td>
</tr>
<tr>
<td>The melodic contours of Romance languages</td>
<td>137</td>
</tr>
</tbody>
</table>
Contents

viii

Inventory 138
Processing prosodic information 141
Prosodic structures in Romance languages 144
Identification of prosodic contours 144
Complex contour 146
Experimental data 148
Sequences of two prosodic words 150
Sequences of three prosodic words 160
Sequences of four prosodic words and more 185
Coordination, enumeration, parenthesis 192
Coordination 192
Enumeration 198
Parenthesis 200
An example of AM prosodic analysis in French 203
An example of ISC prosodic processing in French 208
Conclusion 212

8 Macrosyntax 214
A first approach 215
Three current models for macrosyntax 217
The theory of la lingua in atto 220
Text macrosyntax and prosodic macrosyntax 221
Merging text and intonation 222
Dysfluencies 224
Ponctuants 225
The prosodic eraser 226
Use of dysfluencies 226
Deletions 227
Additions 228
Text and prosodic macrosegments 230
Examples of macrosyntactic analysis 232
French 233
Italian 241
Portuguese 245
Conclusion 248

9 Applications 249
Teaching French prosodic structure 249
Silent reading 252
Eye movement 253
Subvocalization 253
Delta wave synchronization 255

10 Conclusion 256
Quotes from Frédéric Dard (San Antonio) 256

11 WinPitch 259
Sound recording made clear 259
Sound and video 260
Transcription and alignment on the fly 261
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data mining for large speech corpora</td>
<td>262</td>
</tr>
<tr>
<td>Acoustic analysis</td>
<td>266</td>
</tr>
<tr>
<td>Prosodic morphing</td>
<td>270</td>
</tr>
<tr>
<td>Automatic segmentation</td>
<td>270</td>
</tr>
<tr>
<td>Interface with other software</td>
<td>270</td>
</tr>
</tbody>
</table>

| References                                   | 272  |
| Analyzed corpora                              | 285  |
| Author index                                  | 287  |
| Subject index                                 | 290  |
List of figures and maps

### Figures

1.1 Respiration cycle, without phonation (top) and with phonation (bottom)  
   
1.2 An example of an out-of-breath speaker (NS), interrupting the phonation process by pauses longer than usual  
   
1.3 Source-filter model of phonation  
   
1.4 Interactions in the source-filter model between phonation and emotions  

1.5 Extreme cases of the emotion–phonology relationship: emotion dominates phonology (extreme stress or anger), and phonology dominates emotions  

1.6 An example of melodic curve, interrupted at segments without voicing (including pauses and silence), with the fundamental frequency (top), intensity (middle) and wave (bottom) curves  

1.7 Narrowband spectrogram to visualize harmonics corresponding to the fundamental frequency curve  

1.8 Staircase duration curves showing the evolution of syllabic duration  

1.9 Bézier duration curves showing the evolution of syllabic duration  

1.10 Map: Romance dialects  
   
2.1 Rousselot kymograph  

2.2 Measure of the laryngeal period, directly (top), or indirectly from the duration of a packet of ten periods  

2.3 Spectrogram printed on thermo-sensitive paper  

2.4 The ten basic intonation patterns for French by Delattre (1966)  

2.5 An example of analysis with the software Waves™ of the sentence *Jim builds a big daisy-chain* (from ToBI, 1999)  

2.6 An example of a fundamental frequency curve with a wide band spectrogram displayed underneath (from Delais-Roussarie et al., 2015)  

---

x
List of figures and maps

2.7 WinPitch display 26
3.1 Pitch curve using a linear scale in Hz 31
3.2 Melody example of perceived pitch curve transcribed on a musical range for English (Jones, 1909) 33
3.3 Musical transcription used by Fónagy and Magdics (1963) for English 33
3.4 Unstressed syllables, static tones, and contours for English (Armstrong & Ward, 1931) 34
3.5 Unstressed syllables, static tones, and contours for German (von Essen, 1956) 34
3.6 Stressed and final syllables pitch transcribed as static tones in English (Pike, 1945) 34
3.7 Melody contours of groups for English oral (Palmer & Blandford, 1924) 35
3.8 Melody movements by syllable for English (Bolinger, 1961) 35
3.9 Simplified musical range has four levels for French: 1 Low, 2 Average, 3 High, 4 Acute (Léon & Martin, 1969) 35
3.10 Example of comparison of melody movements on all the syllables of statement read by two speakers of the Valley of Aoste 36
3.11 Prosogram for the automatic determination of pitch starting from the syllabic segmentation and of the glissando threshold (Mertens, 2004) 37
3.12 Variations of rising melody contours transcribed with the ToBI notation 39
3.13 In the ToBI transcription, a high tone can correspond to a rising (on the left) or a falling contour (on the right) 40
3.14 An example of an intonative period boundary detection from the four parameters (from Lacheret-Dujour & Victorri, 2002) 42
4.1 An example of metrical grid 47
4.2 Degrees of stress obtained by counting the number of stress nodes 48
4.3 The (revised) Autosegmental-Metrical Prosodic Structure 51
5.1 Micromelodic fundamental frequency dips (circled) due to the presence of a voiced stop [d] in the sequence ... actividades denunciadas por Traffic 60
5.2 Two examples in French (c’est ma maman “this is my mother,” and c’est mon papa “this is my father”) 67
5.3 Variants of modality melodic contours located on the last stressed syllable (declarative case) or the last syllable, stressed or not (interrogative case) 70
5.4 Voulez-vous du thé du café ou du chocolat? 71
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.5</td>
<td>Voulez-vous du thé du café du chocolat?</td>
<td>71</td>
</tr>
<tr>
<td>5.6</td>
<td>Declarative vs. imperative conclusive melodic contours</td>
<td>73</td>
</tr>
<tr>
<td>5.7</td>
<td>Emphasis on a declarative contour gives an imperative contour; emphasis on an interrogative contour gives a surprise contour</td>
<td>74</td>
</tr>
<tr>
<td>5.8</td>
<td>Implicative contour (evidence), a moderate rise followed by a large fall</td>
<td>74</td>
</tr>
<tr>
<td>5.9</td>
<td>Bell shaping on a declarative contour gives an evidence contour, and on an interrogative contour gives a doubt contour</td>
<td>75</td>
</tr>
<tr>
<td>5.10</td>
<td>Interrogative vs. surprise conclusive melodic contours</td>
<td>75</td>
</tr>
<tr>
<td>5.11</td>
<td>Implicative interrogative contour (doubt), a large rise followed by a moderate fall</td>
<td>75</td>
</tr>
<tr>
<td>5.12</td>
<td>A non-planar partial structure [A [B] C], not well-formed for a prosodic structure</td>
<td>83</td>
</tr>
<tr>
<td>5.13</td>
<td>A prosodic structure without connexity, well-formed for a prosodic structure with non-integrated parentheses</td>
<td>84</td>
</tr>
<tr>
<td>5.14</td>
<td>A domain between two consecutive C0 contours, where contours C1 inside the domain must be realized phonetically with enough similarities</td>
<td>85</td>
</tr>
<tr>
<td>5.15</td>
<td>Three basic planar hierarchic configurations</td>
<td>93</td>
</tr>
<tr>
<td>5.16</td>
<td>Three basic prosodic hierarchical configurations</td>
<td>95</td>
</tr>
<tr>
<td>5.17</td>
<td>Prosodic word shortest duration in ms for various speech styles (Martin, 2014b), political discourse, narrative, conference, radio news, university lecture</td>
<td>99</td>
</tr>
<tr>
<td>5.18</td>
<td>A car license plate difficult to read</td>
<td>99</td>
</tr>
<tr>
<td>5.19</td>
<td>A telephone number difficult to read and to remember</td>
<td>99</td>
</tr>
<tr>
<td>5.20</td>
<td>Prosodic word longest duration in ms for various speech styles (Martin, 2014b)</td>
<td>102</td>
</tr>
<tr>
<td>5.21</td>
<td>Two ways to obtain eurhythmicity: balancing the number of syllables or varying the speech rate</td>
<td>103</td>
</tr>
<tr>
<td>5.22</td>
<td>Syllabic duration in function of the number of syllables in prosodic words</td>
<td>104</td>
</tr>
<tr>
<td>5.23</td>
<td>A phrasing non-congruent to syntax, involving a syntactic clash</td>
<td>106</td>
</tr>
<tr>
<td>5.24</td>
<td>Example of EEG spectral analysis (channel 28 or Pz) of evoked potential for a stimulus of a sequence of pure tones with no temporal structure</td>
<td>109</td>
</tr>
<tr>
<td>5.25</td>
<td>Example of EEG spectral analysis (channel 28 or Pz) of evoked potential for a stimulus of a sequence of pure tones with a temporal structure</td>
<td>109</td>
</tr>
<tr>
<td>5.26</td>
<td>EEG Theta waves synchronized by Delta pulses</td>
<td>111</td>
</tr>
<tr>
<td>5.27</td>
<td>The seven syllables constraint is actually a duration constraint, governed by Delta waves</td>
<td>114</td>
</tr>
</tbody>
</table>
5.28 Delta waves synchronize the transfer of chunks of syllables from short-term memory

5.29 The eurhythmicity constraint is linked to the relative stability of consecutive periods of Delta waves

5.30 Chunks of syllables of more than four syllables must be identified in long-term memory

5.31 A hierarchy which contradicts the graphic structure

5.32 A simple enumeration by juxtaposition of the three stress groups

5.33 A structure corresponding to the usual graphic representation 522 4436

7.1 The process of reading schematized

7.2 Marking of stressed melodic contours and boundary tones

7.3 An example of a long enumeration made from a sequence of groups of two prosodic words

7.4 Some examples of terminal conclusive contours in various regional realizations (Turin, Rome, Palermo, Naples, Florence)

7.5 An example of melodic contours sequence, showing the contrast of melodic slope in French

7.6 anticoncezional

7.7 que hay una provocación a la discriminación

7.8 que componen Malásia

7.9 nos estados unidos

7.10 preposición

7.11 An example of prosodic structure

7.12 L’idée était simple

7.13 L’idea era semplice

7.14 La idea era simple

7.15 La idea era simple

7.16 A idea era simples

7.17 Idea era simplà

7.18 In pericolo poi

7.19 cuando se constate

7.20 apelidado de Óscar

7.21 Com les formigues

7.22 La più grande delle piroghe misura cinque metri

7.23 Siguiendo Cn este modelo

7.24 Segundo a especialista

7.25 na cidade de York

7.26 Nascido no Japão

7.27 les garçons de piste

7.28 La plus grande des pirogues mesure cinq mètres
mais les scientifiques japonais
La cuisiner de Sant Pol de Mar
Configuration I, Cx Cx C0
Les romans ont un début et une fin
on rend cette interdiction strictement inefficace
Configuration I, Cx Cx C1
ainsi sa nouvelle gamme de combinés présentés lundi
A saturation of melodic contrasts in the long syntagm
anniversaire de la mort au combat en dix-sept cent dix-huit du roi Charles Douze
Configuration II, Configuration II Cy Cx C0
de se livrer à des affrontements en règle
cette maladie est devenue une pathologie changeante et multiforme
Configuration II, Cy Cx C1
C'est au travers de cette relation qu'il instaurera à ces deux personnes
Configuration III, Cx Cy C0
Certsins de ces bâtiments préfabriqués se sont révélés dangereux
Neuf cents policiers n'ont pu, cependant, empêcher les bagarres recherchées de part et d'autre
Configuration III, Cy Cx C1
A saturation of melodic contrasts in the long syntagm
anniversaire de la mort au combat en dix-sept cent dix-huit du roi Charles Douze
Configuration I, Cx Cx Cc
B, antiparasiit C anticoncezionali
Tromanzi hanno un inizio e una fine
Los romances tienen un inicio y un fin
El romances tenen un inici i un final
Os romances têm um início e um fim
Avião de papel no Espaço
Configuration I, Cx Cx Cc
Um cão de raça terra nova
è permittido matar um escocês
Os vídeos sobre Cn actividades paranormais
La recomendación planta a los Estados
Miçarea separatistā bascā a comis noi atentate
i despés que el vedell ataqués un dels homes que el volia lligar
Configuration, I Cx Cx C1
| 7.62 | *A escolha da carreira profissional* | 176 |
| 7.63 | Configuration II, Cy Cx C0 | 177 |
| 7.64 | in coppie nelle quali il padre è sieropositivo | 177 |
| 7.65 | sara arroolato dai carabinieri e addestrato | 178 |
| 7.66 | probabilmente sfuggito al controllo del padrone | 178 |
| 7.67 | Accasta este o dilemă insolubilă | 179 |
| 7.68 | Configuration II, Cy Cx Cc | 179 |
| 7.69 | Situazione periferica a Portugalie o menține într-o poziție marginală în raport cu fluxurile din Est | 180 |
| 7.70 | i després que el ve* dell* ataqués un dels homes que el volia lliurar | 180 |
| 7.71 | Unele dintre aceste clădiri prefabricate s-au dovedit periculoase | 181 |
| 7.72 | Configuration III, Cy Cx C0 | 181 |
| 7.73 | Romanele Cc au un început C1 și un sfârșit | 182 |
| 7.74 | Alarmă la școala britanică | 182 |
| 7.75 | Configuration III, Cy Cx Cc | 183 |
| 7.76 | che trasferirsi in USA | 183 |
| 7.77 | Poc* s* minuts després de les set de la tarda, el ve* dell* | 184 |
| 7.78 | probabilmente sfuggito al controllo del padrone | 184 |
| 7.79 | Un grup de cercetători germani a rezolvat enigma | 185 |
| 7.80 | les médecins de l’Académie des sciences médicales | 186 |
| 7.81 | El catalán es la ochna y ocho lengua del mundo | 186 |
| 7.82 | L’acadèmia de la llengua catalana l’Institut d’Estudis Catalans IEC | 187 |
| 7.83 | Le programme de recherche a débuté en deux mille deux | 187 |
| 7.84 | Poche zampate per attirare l’attenzione del piantone del comando provinciale dei carabinieri | 188 |
| 7.85 | Una nuova e divertente gymnastica con la palla | 189 |
| 7.86 | es necesario alfabetizar a cuatro millones de personas cada año | 189 |
| 7.87 | La resolución propone que se pongan en marcha dispositivos nacionales de autocontrol | 190 |
| 7.88 | Així abans que acabiel dos mil vuit | 190 |
| 7.89 | Com cerca de sete centímetros de comprimento | 191 |
| 7.90 | In Germania violenza rassistà a depasit limita | 191 |
| 7.91 | je vous suggère d’installer des volets des rideaux et des voilages | 193 |
| 7.92 | je vous conseille d’étudier le néerlandais le danois et le norvégien | 193 |
| 7.93 | *Jamais Barnabé* Jean-David ni Mamadou ne seraient prêts à venir travailler le samedi | 194 |
7.94 *Le muret le donjon et l’église sont de style roman*

7.95 Two possible hierarchical configurations [ABC] and [[A] [B] [C]] for postverbal accentual units, resulting in sequences similar contours rising, rising, falling

7.96 Different groupings coordinated units A, B, and C [ABC] and [[A] [B] [C]] resulting in different sequences of melodic contours, falling, falling, rising or rising, rising, and rising in the preverbal case

7.97 *le vélo le roller ou l’aviron comptent parmi les activités populaires sur le campus*

7.98 *le muret le donjon et l’église sont de style roman* in which the coordinate units are subject of the Verb Phrase *sont de style roman*

7.99 A parallel realization where the first stress groups are coordinated with the conjunction *ni*, associated in each case with an emphatic accent

7.100 *On peut livrer le lavabo la baignoire ou l’évier sans acompte de votre part*

7.101 Enumeration in Italian of numbers (1.46, 1.47, 1.44, 2.78, 2.41), sequence of Cc contours ending each prosodic group, terminated by C0 conclusive on the last item (giorni)

7.102 *…B C2 antiparasiti Ce] [C C2 anticoncezionali Ce] … [M Cc antitlopi C0]* Primo Levi enumeration C2 Cc, C2 Cc … Cc C0

7.103 *Uma e qui pa de cien tas do instituto de reabilitação de Chicago*

7.104 *Il fatto che, in quel mondo, gli uomini*

7.105 *che trasferirsi in USA*

7.106 *permant una utilización más segura de algo que, evidentemente*

7.107 *Ancora i giapponesi, per contro hanno la più alta incidenza mondiale*

7.108 **Le donne giapponesi, per esempio, hanno un’incidenza di tumori alla mammella**

7.109 *Le colèreux garçon ment à sa mère (from Jun & Fougeron, 2002)*

7.110 *le colèreux et mauvais garçon ment à sa mère (from Jun & Fougeron, 2002)*

7.111 A counterexample *le garçon colèreux ment à sa mère (from Jun & Fougeron, 2002)*

7.112 *Ou le donjon ou le minaret ou les murailles doivent être restaurés*
List of figures and maps xvii

7.113 Identification of prosodic events 209
7.114 Classification of prosodic events 209
7.115 Retrieving the prosodic structure 210
8.1 Macrosyntactic analysis of text and intonation 223
8.2 Adding a corrective syntagm after the nucleus: [en Angleterre dans le métro] nucleus 229
8.3 Adding syntactic segment to the prenucleus [mes parents m’emmenaient] 229
8.4 [Je confirme que le premier ministre Elio Di Rupo m’a parlé de cette situation et euh de l’inquiétude des brasseurs] C1 [belges] C1 [par rapport à ce qu’était la consommation française …] 230
8.5 Melodic curve with the pitch movements on stressed syllables circled 234
8.6 The ISC schema of the example Les vieux graphistes ou les anciens je devrais dire graphistes pas les vieux quelquefois lorsqu’ils voient les mises en page de certaines revues ou de certains journaux ils se mettent les mains sur la tête 236
9.1 Variants of conclusive contours: declarative, imperative, implicative, interrogative, surprise, doubt 250
9.2 Three prosodic structures organizing a sequence of three stress groups 251
11.1 WinPitch command, alignment, navigation, and analysis windows 260
11.2 Example of video analysis 261
11.3 Assisted alignment by slowing down speech playback 262
11.4 Automatic IPA transcription from orthographic text and morphological and syntactic labeling 263
11.5 Fine tuning of speech segments limits with the help of a simultaneously displayed spectrogram 263
11.6 Automatic segmentation from spectrographic transitions 264
11.7 Batch processing of a large set of conjunctions obtained from a concordance analyzer with their left and right contexts 265
11.8 Entering the key word “parce que” and selecting a Transcriber file 266
11.9 Table generated automatically listing the occurrences of the entered keyword 267
11.10 Automatic generation of text from alignment files and selection of the entered key word 267
List of figures and maps

11.11 Command window displaying the available pitch tracking algorithms that can be used on any user selected sections of the speech recording 268
11.12 Most common sources of errors for F0 tracking 269
## List of tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Variants of modality</td>
<td>69</td>
</tr>
<tr>
<td>5.2</td>
<td>Phonological description of modality variants using the features +/- Rising, +/- Ample, and +/- Bell shaped</td>
<td>70</td>
</tr>
<tr>
<td>5.3</td>
<td>Phonological description of modality contours</td>
<td>91</td>
</tr>
<tr>
<td>5.4</td>
<td>System of contrasts Cx C0</td>
<td>92</td>
</tr>
<tr>
<td>5.5</td>
<td>System of contrasts Cx C1 C0</td>
<td>93</td>
</tr>
<tr>
<td>5.6</td>
<td>System of contrasts C1 Cx C0</td>
<td>94</td>
</tr>
<tr>
<td>5.7</td>
<td>System of contrasts Cx Cx C1</td>
<td>94</td>
</tr>
<tr>
<td>5.8</td>
<td>System of contrasts Cy Cx C1</td>
<td>95</td>
</tr>
<tr>
<td>5.9</td>
<td>System of contrasts Cx Cy C1</td>
<td>95</td>
</tr>
<tr>
<td>7.1</td>
<td>Processing the prosodic events Cn, C2, C1 and C0 in the example of Figure 7.5</td>
<td>143</td>
</tr>
<tr>
<td>7.2</td>
<td>Static phonological description of romance languages melodic contours</td>
<td>148</td>
</tr>
<tr>
<td>7.3</td>
<td>Configurations of three successive melodic contours</td>
<td>149</td>
</tr>
<tr>
<td>7.4</td>
<td>Percentage of realizations according to the coordination type</td>
<td>195</td>
</tr>
<tr>
<td>7.5</td>
<td>Sequence of prosodic events triggering the storage-concatenation process</td>
<td>211</td>
</tr>
<tr>
<td>8.1</td>
<td>Gars and Lablita equivalence</td>
<td>220</td>
</tr>
<tr>
<td>8.2</td>
<td>Utterance segmentation into prosodic words indicating the melodic contours on the last stressed syllable and the type of contour in the process of storage-concatenation</td>
<td>236</td>
</tr>
</tbody>
</table>
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 minguograph (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;
Acknowledgments

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.