

Cambridge University Press  
0521530695 - Introducing Speech and Language Processing  
John Coleman  
Frontmatter  
[More information](#)

---

## Introducing Speech and Language Processing

This major new textbook provides a clearly written, concise and accessible introduction to speech and language processing. Assuming knowledge of only the very basics of linguistics and written specifically for students with no technical background, it is the perfect starting point for anyone beginning to study the discipline. Students are introduced to topics such as digital signal processing, speech analysis and synthesis, finite-state machines, automatic speech recognition, parsing and probabilistic grammars, and are shown from a very elementary level how to work with two programming languages, C and Prolog. The accompanying CD-ROM contains all the software described in the book, along with a C compiler, Prolog interpreter and sound file editor, thus providing a self-contained, one-stop resource for the learner. Setting a firm grounding in speech and language processing and an invaluable foundation for further study, *Introducing Speech and Language Processing* is set to become the leading introduction to the field.

JOHN COLEMAN is Director of the Phonetics Laboratory and Reader in Phonetics at Oxford University, having previously worked on speech technology at the universities of Hull and York, and AT&T Bell Laboratories. His previous books are *Phonological Representations* (Cambridge University Press, 1998) and *Acoustics of American English Speech* (Springer-Verlag, 1993).

Cambridge University Press  
0521530695 - Introducing Speech and Language Processing  
John Coleman  
Frontmatter  
[More information](#)

---

**Cambridge Introductions to Language and Linguistics**

This new textbook series provides students and their teachers with accessible introductions to the major subjects encountered within the study of language and linguistics. Assuming no prior knowledge of the subject, each book is written and designed for ease of use in the classroom or seminar, and is ideal for adoption on a modular course as the core recommended textbook. Each book offers the ideal introductory material for each subject, presenting students with an overview of the main topics encountered in their course, and features a glossary of useful terms, chapter previews and summaries, suggestions for further reading, and helpful exercises. Each book is accompanied by a supporting website.

Books published in the series  
*Introducing Phonology* David Odden  
*Introducing Speech and Language Processing* John Coleman

Forthcoming:  
*Introducing Phonetic Science* Michael Ashby and John Maidment  
*Introducing Sociolinguistics* Miriam Meyerhoff  
*Introducing Morphology* Maggie Tallerman and S. J. Hannahs  
*Introducing Historical Linguistics* Brian Joseph  
*Introducing Second Language Acquisition* Muriel Saville-Troike  
*Introducing Language* Bert Vaux

Cambridge University Press  
0521530695 - Introducing Speech and Language Processing  
John Coleman  
Frontmatter  
[More information](#)

# Introducing Speech and Language Processing

JOHN COLEMAN  
*University of Oxford  
Phonetics Laboratory*



ISBN 0 521 82365 X hardback  
ISBN 0 521 53069 5 paperback

# Contents

<i>Acknowledgements and copyright notices</i>	x
<i>CD-Rom and companion website</i>	x
<b>1</b> <b>Introduction</b>	<b>1</b>
1.1    About this book	2
1.2    Purpose of this book	2
1.3    Some reasons to use this book	3
1.4    What's in the book (and what's not)	5
1.5    Computational set-up needed for this book	8
1.6    Computational skills that are necessary in order to use the book	9
1.7    Free software suggestions	10
1.8    Book structure	10
<b>2</b> <b>Sounds and numbers</b>	<b>13</b>
2.1    Preparatory assignments	14
2.2    Solutions	21
2.3    Sampling	24
2.4    Quantization	24
2.5    The sampling theorem	27
2.6    Generating a signal	29
2.7    Numeric data types	31
2.8    The program	34
2.9    Structure of a loop	35
2.10    Structure of an array	37
2.11    Calculating the cosine values	38
2.12    Structure of the program	39
2.13    Writing the signal to a file	40
<i>Chapter summary</i>	43
<i>Further Exercises</i>	43
<i>Further reading</i>	46
<b>3</b> <b>Digital filters and resonators</b>	<b>47</b>
3.1    Operations on sequences of numbers	48
3.2    A program for calculating RMS amplitude	48
3.3    Filtering	50
3.4    A program for calculating running means of 4	52
3.5    Smoothing over a longer time-window	54
3.6    Avoiding the need for long windows	54
3.7    IIR filters in C	61

3.8	Structure of the Klatt formant synthesizer	62
	<i>Chapter summary</i>	68
	<i>Exercises</i>	68
	<i>Further reading</i>	69
4	Frequency analysis and linear predictive coding	71
4.1	Spectral analysis	72
4.2	Spectral analysis in C	72
4.3	Cepstral analysis	79
4.4	Computation of the cepstrum in C	80
4.5	Pitch tracking using cepstral analysis	83
4.6	Voicing detection	86
4.7	$f_0$ estimation by the autocorrelation method	90
4.8	Linear predictive coding	95
4.9	C programs for LPC analysis and resynthesis	100
4.10	Trying it out	106
4.11	Applications of LPC	106
	<i>Chapter Summary</i>	109
	<i>Further exercises</i>	109
	<i>Further reading</i>	110
5	Finite-state machines	111
5.1	Some simple examples	112
5.2	A more serious example	113
5.3	Deterministic and non-deterministic automata	116
5.4	Implementation in Prolog	118
5.5	Prolog's processing strategy and the treatment of variables	129
5.6	Generating strings	132
5.7	Three possibly useful applications of that idea	134
5.8	Another approach to describing finite-state machines	135
5.9	Self-loops	137
5.10	Finite-state transducers (FSTs)	139
5.11	Using finite-state transducers to relate speech to phonemes	144
5.12	Finite-state phonology	149
5.13	Finite-state syntactic processing	153
	<i>Chapter summary</i>	156
	<i>Further exercises</i>	156
	<i>Further reading</i>	156
6	Introduction to speech recognition techniques	157
6.1	Architectures for speech recognition	158
6.2	The pattern-recognition approach	166
6.3	Dynamic time warping	168
6.4	Applications	177
6.5	Sources of variability in speech	181
	<i>Chapter summary</i>	182
	<i>Further reading</i>	183

<b>7</b>	<b>Probabilistic finite-state models</b>	<b>185</b>
7.1	Introduction	186
7.2	Indeterminacy: <i>n</i> -gram models for part-of-speech tagging	187
7.3	Some probability theory for language modelling	190
7.4	Markov models	192
7.5	Trigram models	198
7.6	Incompleteness of the training corpus	202
7.7	Part-of-speech model calculations	209
7.8	Using HMMs for speech recognition	210
7.9	Chomsky's objections to Markov models and some rejoinders	213
	<i>Chapter summary</i>	219
	<i>Further reading</i>	219
<b>8</b>	<b>Parsing</b>	<b>221</b>
8.1	Introduction	222
8.2	A demo	222
8.3	'Intuitive' parsing	223
8.4	Recursive descent parsing	225
8.5	The simplest parsing program	232
8.6	Difference lists	233
8.7	Generating a parse tree	236
8.8	Syllabification	238
8.9	Other parsing algorithms	242
8.10	Chart parsing	242
8.11	Depth-first vs. breadth-first search	245
8.12	Deterministic parsing, Marcus parsing and minimal commitment parsing	246
8.13	Parallel parsing	249
	<i>Chapter summary</i>	249
	<i>Further reading</i>	250
<b>9</b>	<b>Using probabilistic grammars</b>	<b>251</b>
9.1	Motivations	252
9.2	Probabilistic context-free grammars	256
9.3	Estimation of rule probabilities	258
9.4	A practical example	261
9.5	A limitation of probabilistic context-free grammars	267
9.6	Tree adjoining grammars	268
9.7	Data-oriented parsing	271
	<i>Chapter Summary</i>	272
	<i>Conclusion and suggestions for further reading</i>	272
	<b>Appendix: The American Standard Code for Information Interchange (ASCII)</b>	<b>275</b>
	<i>Glossary</i>	277
	<i>References</i>	293
	<i>Index</i>	299

# Acknowledgements and copyright notices

The production of a textbook such as this is very much a team effort involving many people besides the author. I reserve my greatest thanks for the students who have followed my introductory courses on speech and language processing during the last ten years. Their interest in and enthusiasm for the subject sustained my efforts to put these lectures into publishable form, and their critical comments and questions led to many improvements. I would also like to thank my secretary, Celia Glyn, who patiently transcribed an early set of recordings of the classes, thus laying down the foundation of the text. Equally, my ability to develop these course materials, and even write the textbook, is due to the technical support of Andrew Slater and other computing support staff in my department. In particular, the DTW program in chapter 6 is abbreviated from code developed by Andrew, in relation to research we have conducted together (Slater and Coleman 1996).

From Cambridge University Press, I was encouraged first by Christine Bartels, then by Andrew Winnard and Juliet Davis-Berry to get on and finish it! The reviewers that Cambridge engaged to comment in detail on the manuscript deserve praise as well as thanks for their thorough critical commentaries. Their suggestions and complaints were most valuable. In addition, Robin Coleman provided valuable help in compiling the index.

Much of the material on the CD-ROM is copyright and is provided for use under licence only

by the owner of the book. Some of the material was written by others, and is only included here with their permission under terms stated in the relevant files on the disk. For classroom use, multiple copies of the software may be made with permission from the Copyright Clearance Centre. Otherwise, copying and distribution of any of the material on the CD-ROM is prohibited.

For the non-copyright intellectual property of all the methods of speech and language processing that are explained in the following pages, I trust that the citations of prior work will be acceptable indicators of my indebtedness to earlier workers in the field, with respect to whom we are all students.

## CD-ROM and companion website

The CD-ROM included with this book contains all the software used in the course, including:

- Two programming languages (C and Prolog)
- A sound file editor and a text file editor
- All of the program listings printed in the text, and a few extras
- Licence details

Neither Cambridge University Press nor John Coleman make any warranties, express or implied, that the programs contained in this book and on the CD-ROM and companion website are free of error, or are consistent with any particular standard of merchantability, or that they will meet your requirements for any particular application. They should not be relied on for solving a problem whose incorrect solution could



result in injury to a person or loss of property. If you do use the programs in such a manner, it is at your own risk. The authors and publisher disclaim all liability for direct or consequential damages resulting from your use of the program.

I may improve the software in future, as may other developers of the software included on the disk. Consequently, revised versions of the software, or links to revised or alternative resources, may be made available through the companion website, [www.islp.org.uk](http://www.islp.org.uk). Information about

bugs in the software that may be found after publication will also be placed there, so if any of the programs appear to be defective, please check there for a possible upgrade before attempting to contact me. (You can send email to me by visiting the companion website.)

Because material on the internet changes so frequently, there are no URLs (i.e. addresses of web pages) in the text, except for [www.islp.org.uk](http://www.islp.org.uk). Instead, I have provided some links to useful websites at [www.islp.org.uk](http://www.islp.org.uk).