

Language, Technology, and Society

Richard Sproat

(Oregon Health & Science University)

Oxford: Oxford University Press, 2010, xiii+286 pp; hardbound,
ISBN 978-0-19-954938-2, £25.00

Reviewed by

Francis Bond

Nanyang Technological University

Sproat's book is an enjoyable collection of loosely connected discussions on writing systems and their decipherment, ancient and modern speech technology, and machine translation. The book arose from a course on Language, Technology, and Society but it's definitely not a textbook—instead of broad coverage, it offers a few enjoyable in-depth discussions. In fact, it feels somewhat like a coffee-table book for computational linguists.

The book starts off by showing how *language* and *tool creation and use* are both peculiar characteristics of human beings. This leads nicely into the theme of language technology as a combination of these distinctive human traits. Chapter 2 then presents writing as the first case of language technology, contrasting writing with other symbolic systems.

Chapter 3 gives a detailed discussion of how writing represents language. The main point is that writing systems all end up basically representing sound. This is true even though writing appears to have been independently invented in at least three places, and has a huge variety of forms, with many differences in script type, glyph combination, and direction. Sproat argues convincingly, with examples from many different languages, that sound is the level represented because it is only at this level that there are few enough units to be remembered—syllables, words, concepts, or phrases are too numerous.

Sproat next discusses the decipherment of writing systems where the mapping from symbol to sound has been lost. Sproat starts with the classic case of hieroglyphics, deciphered with the help of the parallel text on the Rosetta stone. He then discusses potential scripts with fewer extant samples, emphasizing the point that we may never be able to decipher some of them—without some gold standard we can never know whether the mapping produced is actually correct. Indeed, for those sets of symbols for which we have limited data and no linking to other texts, we may never know whether they are even writing systems.

The following chapter talks about writing, literacy, and technology. Sproat shows that complexity of a script appears to have little to do with literacy—general measures of well being (such as health and schooling) are much more important. The chapter ends with a very speculative discussion of the danger of losing the permanence of writing (as in Orwell's *Nineteen Eighty-Four*) and returning to a culture where we have no reliable records.

Chapter 6 gives a brief history of mechanized speech and language technology. It begins with a detailed description of early attempts at speech synthesis and typewriters, and then discusses telegraphy, computer encoding of text, and Braille. Chapter 7 then moves to an overview of modern speech technology including both recognition and

synthesis. This section presents the state of the art without losing the reader in technical details. It is very clearly written, well referenced, and authoritative.

Chapter 8 is mainly about machine translation. It combines a balanced historical discussion of research and development in machine translation with a very accessible introduction to statistical machine translation. It is a pity that the same level of detail couldn't have been extended to other fields of natural language processing, such as information retrieval and question answering.

The book finishes with a discussion of the future of language technology. The first section of the closing chapter predicts how language technology, especially speech technology, will improve in the near future. This section includes a good discussion of the importance of investment in developing technology—not everything that could be done will be; the greatest improvement will be seen in those areas in which people are prepared to invest time and money. It was nice to see this often neglected topic addressed. The final section talks about the social implications of language technology, warning that they are not all positive.

Overall the book is a pleasure to read,¹ with well-chosen illustrations, nice examples, illuminating illustrations, and numerous examples of different writing systems. The author is clearly knowledgeable about writing systems, especially the decipherment of scripts with little extant text, as well as speech processing, his main field of research. These sections are very clearly written, well referenced, and authoritative. The sections on the social implications are more speculative.

The book's weakness is that the coverage is very idiosyncratic. For example, there is almost no discussion of printing, which is surely one of the great language technologies. In contrast, several fairly minor topics, such as Blissymbolics and the Phaistos disc, are treated very thoroughly. In particular, there is no discussion on the bandwidth and compressibility of various scripts and almost nothing on computer-mediated communication and its effect on language. For a book on language and technology published in 2010, this is a disappointment. The structure is also a little inconsistent: Chapters 3 and 6 have summaries, Chapter 7 has a synopsis, and the rest just finish.

In summary, Sproat's book is an enjoyable collection of essays on language, technology, and society, but still leaves room for a more definitive tome.

Francis Bond is an Associate Professor in the Division of Linguistics and Multilingual Studies, Nanyang Technological University, Singapore. His main research interests are in natural language understanding, machine translation, HPSC, and WordNets. Previously he was a researcher at Nippon Telegraph and Telephone (NTT) and the National Institute of Information and Communications Technology, Japan (NICT). E-mail: bond@ieee.org.

1 Except for the fact that it uses endnotes, rather than footnotes, which means you have to have jump backwards and forwards to read the otherwise excellent notes.