Lexical Segmentation in Spoken Word Recognition

Matthew Harold Davis

Birkbeck College,

University of London

Thesis submitted for the degree of Doctor of Philosophy

March 2000

Abstract

This thesis examines an important issue in spoken word recognition; how the perceptual system segments connected speech into lexical units or words. Research on this topic has investigated the role of different sources of information in dividing up the speech stream: acoustic cues in the speech signal, statistical regularities in the structure of the language or through the identification of individual lexical items.

This research focuses on cases in which the location of word boundaries may be ambiguous by one or more of these segmentation mechanisms using words embedded at the onset of longer words (such as *cap* in *captain*). The ambiguities proposed for onset-embedded words have motivated accounts of segmentation based on competition between alternative parses of speech into words. In these accounts, the recognition of embedded words is delayed until after their offset when subsequent input rules out longer competitors. In this thesis it is demonstrated that training a simple recurrent network to activate a representation of all the words in a sequence allows a connectionist network to learn the appropriate delay to allow the identification of onset-embedded words without requiring directly implemented competition between words.

Both lexical competition and recurrent network models assume ambiguity between onset-embedded words and equivalent syllables in longer competitors. Acoustic analysis carried out in this thesis confirms the presence of reliable acoustic differences between syllables in short and long words. A series of experiments using gating and cross-modal priming suggest that the perceptual system uses these acoustic differences to discriminate embedded words from the onset of longer competitors and that match or mismatch with longer competitors may be less important for the identification of onset-embedded words. These results are interpreted within a revised version of the recurrent network model, incorporating input representing the acoustic differences between syllables in short and long words.

Acknowledgements

I am extremely grateful to my supervisors, William Marslen-Wilson and Gareth Gaskell, for their guidance and support throughout the course of this work. I have also benefited from discussions with John Bullinaria, Morten Christiansen, Gary Cottrell, Tom Loucas and Billi Randall on many of the topics contained in this thesis.

At a physical rather than intellectual level, this thesis could not have been written without the assistance of Dr. Whelan, Mr. Cobb and many others at the Middlesex hospital. The care and support of friends and family during my illness helped make this time a frustration rather than a torment.

A further debt of thanks is also owed to my colleagues past and present in the Centre for Speech and Language both at Birkbeck and in Cambridge who provided such a friendly atmosphere in which to work. In particular, thanks must go to Billi Randall and Tom Loucas who helped make my time at CSL so enjoyable. Thanks are also due to Lolly Tyler who faced my divided loyalties in the past year with great stoicism and to Helen Moss and Kathy Rastle for help with proof-reading.

I would like to thank Joe Devlin, Mike Ford and Richard Russell for hospitality and barbecues in Cambridge and Liz Gresham for assisting with matters of style. I also wish to thank my parents for their encouragement throughout my studies. I have kept them from retirement for more years than I should have.

Most important of all though, I would like to thank Maggie Kemmner for advice and encouragement throughout and for the love and support that helped me through the most difficult times.

This research was supported by a studentship from the Engineering and Physical Sciences Research Council.

Contents

List of tables

List of figures


HEY HOBBES, WANT TO SEE AN ANTELOPE?
AN ANTELOPE?!
C'MON!
SEE, SHE'S COMING DOWN THE LADDER TO HER BOYFRIEND'S CAR!
YOU'RE NOT LAUGHING.
IT'S NOT FUNNY.
WATTERSON
7.2