## Contents

**Foreword**

1 General introduction

1.1 Target-undershoot in speech production
   1.1.1 The classical model of vowel target-undershoot
   1.1.2 Interpretations of the target-undershoot model
   1.1.3 Is undershoot the result of articulatory limitations or is it planned?
      1.1.3.1 Input-driven versus output-driven control of articulation
      1.1.3.2 Testing the target-undershoot model

1.2 Perceptual-overshoot and dynamic-specification in vowel identification
   1.2.1 Dynamic-specification versus elaborate target models of vowel perception
   1.2.2 Evidence pro and contra dynamic-specification
   1.2.3 Distinguishing models of vowel perception

2 Formant frequencies of Dutch vowels in a text, read at normal and fast rate

   Introduction
   2.1 Methods
      2.1.1 Speech material
      2.1.2 Segmentation
      2.1.3 Vowels used
      2.1.4 Spectral Analysis
   2.2 Results
      2.2.1 Median values
      2.2.2 Consistency
      2.2.3 Pairwise changes in formant frequencies and duration
      2.2.4 Correlation between formant frequency and duration
      2.2.5 Influence of phoneme context
      2.2.6 Influence of stress
   2.3 Discussion
      2.3.1 Differences between speaking rates: Duration
2.3.2 Differences between speaking rates: Formant frequencies 36
2.3.3 Differences between measuring methods 37
2.4 Conclusions 38

3 Formant movements of Dutch vowels in a text, read at normal and fast rate

Introduction 40
3.1 Methods 41
  3.1.1 Speech material and segmentation 41
  3.1.2 Vowels used 42
  3.1.3 Spectral analysis and formant track sampling method 43
3.2 Results 43
  3.2.1 Duration 43
  3.2.2 Effects of speaking rate on formant frequencies 44
  3.2.3 Correlation between speaking rates 45
  3.2.4 Effects of duration on formant frequencies 46
  3.2.5 Effects of context 48
  3.2.6 Effects of stress 48
3.3 Discussion 49
  3.3.1 Effects of speaking rate 49
  3.3.2 Effects of duration on formant tracks 50
  3.3.3 Effects of context and stress 50
3.4 Conclusions 51

4 The influence of speaking rate on vowel formant track shape as modeled by Legendre polynomials

Introduction 54
4.1 Methods 55
  4.1.4 Measuring differences between formant tracks 55
4.2 Results 57
  4.2.1 Goodness of fit 58
  4.2.2 Legendre polynomial coefficients and their interpretation 59
  4.2.3 Relations between polynomial components 61
  4.2.4 Effects of speaking rate 63
5 The influence of formant track shape on the identification of synthetic vowels

Introduction

5.1 Methods

5.1.1 Isolated vowels

5.1.1.1 Token synthesis
5.1.1.2 Token construction
5.1.1.3 Presentation
5.1.1.4 Subjects

5.1.2 Presentation in synthetic CVC syllables

5.1.2.1 Consonants
5.1.2.2 Vowel segments and syllable construction
5.1.2.3 Presentation and subjects

5.2 Results

5.2.1 Isolated vowel presentation

5.2.1.1 Effects of duration on tokens with level formant tracks
5.2.1.2 Effects of extreme formant excursion sizes on token identification
5.2.1.3 Effects of realistic formant excursion sizes on token identification

5.2.2 Presentation of vowels in context

5.2.2.1 Consistency in responses to synthetic vowels
5.2.2.2 The responses to synthetic consonants and their influence on vowel identification
5.2.2.3 The influence of formant excursion size on vowel identification

5.3 Discussion

5.3.1 The effects of duration
5.3.2 The effects of formant excursion size
5.3.3 The effects of context 91
5.3.4 Relevance for natural speech 92
5.4 Conclusions 93

6  Vowel perception: A closer look at the literature 95

Introduction 96
6.1 An evaluation of the relevant literature 97
   6.1.1 Information present in formant dynamics 97
   6.1.2 Natural versus synthetic speech 98
   6.1.3 Experiments using synthetic speech 98
      6.1.3.1 The paper of Lindblom and Studdert-Kennedy (1967) 99
      6.1.3.2 The paper of Nearey (1989) 102
      6.1.3.3 The paper of Di Benedetto (1989b) 104
      6.1.3.4 The paper of Fox (1989) 105
      6.1.3.5 The paper of Akagi (1993) 107
      6.1.3.6 What factor could induce perceptual-overshoot? 107
   6.1.4 Experiments using natural speech 108
      6.1.4.1 The influence of context on vowel intelligibility 108
      6.1.4.2 The importance of the transition for vowel recognition 109

6.2 Integration of the available results 111
6.3 Conclusions 113

7  General discussion 115

Introduction 116
7.1 Target-undershoot in production 116
   7.1.1 Quasi-stationary formant analysis might give inaccurate values 117
   7.1.2 Too small a difference between normal- and fast-rate speech 118
   7.1.3 A ceiling (floor) in undershoot was already reached 119
   7.1.4 Variation in context has averaged out any difference between speaking rates 120
   7.1.5 Coarticulation was not strong enough to require extra undershoot 122
   7.1.6 Alternative articulation strategies 123
   7.1.7. Does duration control vowel target-undershoot? 123
7.2 Perceptual-overshoot, dynamic-specification, and target models of perception
   7.2.1 Recapitulation of our vowel identification results
   7.2.2 Results from the literature
7.3 Target-undershoot and vowel perception
7.4 Conclusions
7.5 Suggestions for future research

References

Acknowledgements

Summary

Samenvatting (summary in Dutch)

Appendices

A Automatic slope measurement on formant tracks
B Calculating Legendre polynomial coefficients
C Annotated texts with accent transcription
D Formant values and excursion sizes

Name Index

Subject Index