HANDBOOK OF MULTIMODAL AND SPOKEN DIALOGUE SYSTEMS

Resources, Terminology and Product Evaluation

Gibbon/Mertins/Moore (Eds.)

FT MEADE
GenColl
QA 76
.9
.N38
H362
2000
COPY 1

KAP

HANDBOOK OF MULTIMODAL AND SPOKEN DIALOGUE SYSTEMS

Resources, Terminology and Product Evaluation

Dafydd Gibbon
Inge Mertins
Roger K. Moore
(Eds.)

CD ROM Included

HANDBOOK OF MULTIMODAL AND SPOKEN DIALOGUE SYSTEMS

Resources, Terminology and Product Evaluation

THE KLUWER INTERNATIONAL SERIES
IN ENGINEERING AND COMPUTER SCIENCE

HANDBOOK OF MULTIMODAL AND SPOKEN DIALOGUE SYSTEMS

Resources, Terminology and Product Evaluation

edited by

Dafydd Gibbon
University of Bielefeld

Inge Mertins
University of Bielefeld

Roger K. Moore
DERA and 20/20 Speech Limited

Kluwer Academic Publishers
Boston/Dordrecht/London

LIBRARY OF CONGRESS
AUG 29 2000
CIP

Distributors for North, Central and South America:
Kluwer Academic Publishers
101 Philip Drive
Assinippi Park
Norwell, Massachusetts 02061 USA
Telephone (781) 871-6600
Fax (781) 871-6528
E-Mail <kluwer@wkap.com>

Distributors for all other countries:
Kluwer Academic Publishers Group
Distribution Centre
Post Office Box 322
3300 AH Dordrecht, THE NETHERLANDS
Telephone 31 78 6392 392
Fax 31 78 6546 474
E-Mail services@wkap.nl>

Electronic Services <http://www.wkap.nl>

Library of Congress Cataloging-in-Publication

Handbook of multimodal and spoken dialogue systems : resources, terminology, and product evaluation / edited by Dafydd Gibbon, Inge Mertins, Roger K. Moore.
p. cm. -- (Kluwer international series in engineering and computer science ; SECS 565)
Includes bibliographical references and index.
ISBN 0-7923-7904-7 (alk. paper)
1. Natural language processing (Computer science) 2. Automatic speech recognition.
I. Gibbon, Dafydd. II. Mertins, Inge, 1964- III. Moore, Roger, 1952- IV. Series.

QA76.9.N38 H362 2000
006.3'5--dc21

00-044780

Copyright © 2000 by Kluwer Academic Publishers.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, mechanical, photo-copying, recording, or otherwise, without the prior written permission of the publisher, Kluwer Academic Publishers, 101 Philip Drive, Assinippi Park, Norwell, Massachusetts 02061

Printed on acid-free paper.

Printed in the United States of America

LIBRARY OF CONGRESS
AUG 29 2000
CIP

Distributors for North, Central and South America:
Kluwer Academic Publishers
101 Philip Drive
Assinippi Park
Norwell, Massachusetts 02061 USA
Telephone (781) 871-6600
Fax (781) 871-6528
E-Mail <kluwer@wkap.com>

Distributors for all other countries:
Kluwer Academic Publishers Group
Distribution Centre
Post Office Box 322
3300 AH Dordrecht, THE NETHERLANDS
Telephone 31 78 6392 392
Fax 31 78 6546 474
E-Mail services@wkap.nl>

Electronic Services <http://www.wkap.nl>

Library of Congress Cataloging-in-Publication

Handbook of multimodal and spoken dialogue systems : resources, terminology, and product evaluation / edited by Dafydd Gibbon, Inge Mertins, Roger K. Moore.
p. cm. -- (Kluwer international series in engineering and computer science ; SECS 565)
Includes bibliographical references and index.
ISBN 0-7923-7904-7 (alk. paper)
1. Natural language processing (Computer science) 2. Automatic speech recognition.
I. Gibbon, Dafydd. II. Mertins, Inge, 1964- III. Moore, Roger, 1952- IV. Series.

QA76.9.N38 H362 2000
006.3'5--dc21

00-044780

Copyright © 2000 by Kluwer Academic Publishers.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, mechanical, photo-copying, recording, or otherwise, without the prior written permission of the publisher, Kluwer Academic Publishers, 101 Philip Drive, Assinippi Park, Norwell, Massachusetts 02061

Printed on acid-free paper.

Printed in the United States of America

Dedicated to the memory of our colleague, co–author and friend

Christian Benoît

Contents

Editorial Preface xvii

1 Representation and annotation of dialogue 1
1.1 Introduction . . . 1
1.1.1 Goals . . . 1
1.1.2 What is meant by 'Integrated Resources'? . . . 2
1.1.3 Limitations . . . 3
1.2 A preliminary classification of dialogue corpora . . . 5
1.2.1 Dialogue acts . . . 6
1.2.2 Towards a dialogue typology . . . 6
1.3 General coding issues . . . 11
1.4 Orthography . . . 12
1.4.1 Orthographic representation . . . 12
1.4.2 Recommendations . . . 24
1.5 Morphosyntax . . . 26
1.5.1 Morphosyntactic (POS) annotation . . . 26
1.5.2 Recommendations . . . 32
1.6 Syntax . . . 32
1.6.1 Syntactic annotation . . . 32
1.6.2 Recommendations . . . 39
1.7 Prosody . . . 39
1.7.1 Prosodic annotation . . . 39
1.7.2 Recommendations . . . 53
1.8 Pragmatics . . . 54
1.8.1 Pragmatic annotation: functional dialogue annotation . . . 54
1.8.2 Recommendations . . . 66
Appendix A: TEI paralinguistic features . . . 67
Appendix B: TEI P3 DTD: base tag set for transcribed speech . . . 68
Appendix C: A few relevant web links . . . 70
Appendix D: Specimen Annotated Dialogue . . . 70
D.1: Orthographic Transcription . . . 71
D.2: Morphosyntactic annotation . . . 72
D.3: Syntactic annotation . . . 73
D.4: Prosodic Annotation . . . 75
D.5: Pragmatic (Dialogue Act) Annotation . . . 84
D.6: Combined Multi-level Annotation . . . 87
Appendix E: Morphosyntactic annotation of corpora . . . 89
Appendix E.1: English tagset . . . 89
Appendix E.2: Italian DMI codes . . . 95

2 Audio-visual and multimodal speech-based systems 102
2.1 Introduction . . . 102
2.1.1 Terminology . . . 103
2.1.2 Chapter outline . . . 106
2.1.3 Benefits of multimodal systems . . . 106

2.1.4 Input modalities associated with speech 109
2.1.5 Output modalities associated with speech 112
2.1.6 Taxonomies of multimodal applications 114
2.2 Survey of multimodal systems 118
2.3 Evaluation of multimodal systems 122
2.3.1 Types of evaluation 123
2.3.2 Evaluation methodologies 124
2.3.3 Specific evaluation issues 127
2.3.4 Recommendations 129
2.4 Speech input with facial information (audio-visual speech recognition) 129
2.4.1 Face recognition 129
2.4.2 Locating and tracking of other facial features 130
2.4.3 Automatic lipreading systems 131
2.4.4 Integration of audio and visual signals 131
2.5 Speech output with talking heads 132
2.5.1 Control techniques 132
2.5.2 Lip shape computation 137
2.5.3 Talking heads: audio and video output synchronisation . 138
2.6 Speech input with modalities other than faces 138
2.6.1 Recognition of non-speech input modalities 139
2.6.2 Integration in multimodal applications 140
2.7 Speech output in multimedia systems 145
2.7.1 Taxonomy of output modalities 146
2.7.2 Output devices 146
2.7.3 Theoretical issues 147
2.7.4 Summary of recommendations 155
2.8 Technology of multimodal system components 157
2.8.1 Techniques related to face recognition systems 157
2.8.2 Synthesis module 163
2.8.3 Facial models 164
2.8.4 Building conversational agents 173
2.8.5 On-line character and handwriting recognition 178
2.8.6 Gesture recognition 183
2.8.7 Technical issues 190
2.9 Standards and resources for multimodal/multimedia systems . . 190
2.9.1 Standards and resources for monomodal processing 190
2.9.2 Towards standards for multimedia systems 191
2.9.3 Towards standards for hypermedia systems 193
2.9.4 Architectures and toolkits for multimodal integration . . 193
2.9.5 Notational systems 195
2.9.6 Face and audio databases 196

3 Consumer off-the-shelf (COTS) product and service evaluation 204
3.1 Introduction 204
3.1.1 Purpose and scope of this chapter 204
3.1.2 Introduction to speech technologies and classification . . . 204
3.1.3 Automatic speech recognition 205

Google Exhibit 1008
Page 0011

3.1.4 Text-to-speech and speech synthesis 206
3.1.5 Speaker recognition and verification 208
3.1.6 Speech understanding 208
3.1.7 Dialogue control 209
3.2 General remarks 209
3.2.1 Assessment methodology 209
3.2.2 Subjective assessment measures 213
3.2.3 Acoustic environment 214
3.2.4 Comparing several systems 216
3.3 Command and control systems 216
3.3.1 Typical systems 216
3.3.2 Typical issues 218
3.3.3 Evaluation design 220
3.3.4 Examples 222
3.4 Document generation 227
3.4.1 Typical systems 227
3.4.2 Typical issues 228
3.4.3 Evaluation design 229
3.4.4 Examples 229
3.5 Services and telephone applications 233
3.5.1 Typical systems 233
3.5.2 Typical issues 234
3.5.3 Evaluation design 234
3.5.4 Examples 235
3.6 Conclusion and summary of recommendations 238

4 Terminology for spoken language systems 240
4.1 Introduction 240
4.1.1 Terminology standards 240
4.1.2 Termbank users 242
4.1.3 Chapter outline 243
4.2 Terminological basics 243
4.2.1 Central notions in terminological theory 243
4.2.2 Relations between terms 247
4.3 The organisation of terminology 249
4.3.1 The onomasiological and semasiological perspectives . . . 249
4.3.2 Terminological macrostructures and microstructures . . . 251
4.4 Spoken Language terminology 252
4.4.1 The hybrid character of SL terminology 252
4.4.2 Toward a microstructure for SL terminology 253
4.4.3 Recommendations on termbank development 259
4.4.4 Recommendations for further reading 260
4.5 Relational databases 261
4.5.1 Components of a relational database 261
4.5.2 Structures in the relational model 261
4.5.3 Codd's definition of a relational database system 262
4.5.4 Query language 262
4.5.5 Software implementations 262

4.5.6 Distribution of data generation over time 263
4.5.7 Distribution of data generation over resources 263
4.5.8 Required system components 264
4.6 Terminology Management Systems (TMSs), databases, and interchange formats 264
4.6.1 MultiTerm 264
4.6.2 ITU Telecommunication Terminology Database: TERMITE 265
4.6.3 TERMIUM – Canadian Linguistic Data Bank 267
4.6.4 EURODICAUTOM 268
4.6.5 MARTIF terminology interchange format (ISO 12200) . . 269
4.7 The EAGLET Term Database: an SL termbank 271
4.7.1 A hypergraph–based approach 271
4.7.2 Conceptual parts 272
4.7.3 Information storage 272
4.7.4 System components 272
4.7.5 Structure 273
4.7.6 EAGLET macrostructure for SL terminology 273
4.7.7 EAGLET microstructure for SL terminology 275
4.7.8 Using the EAGLET Term Database 277
4.7.9 Future work 280

5 Reference materials 281
5.1 Introduction 281
5.2 Organisations and infrastructure 282
5.2.1 Speech resources, agencies, and associations 282
5.2.2 Archives, general information 291
5.2.3 Education and conferences 293
5.3 "SLP at Work" 296
5.3.1 Speech interfaces 296
5.3.2 Telecommunications and broadcast 297
5.3.3 New services 298
5.3.4 SLP as a research tool 298
5.4 SLP procedures, tools, and formats 301
5.4.1 Annotation 302
5.4.2 Validation, evaluation 303
5.4.3 Tools and standards 304
5.4.4 Text 308
5.5 Technology 309
5.5.1 Alphabets 310
5.5.2 Networks 310
5.5.3 File formats 322
5.5.4 Programming 324
5.5.5 Storage 326

Bibliographical references 329

A SAMPA and X-SAMPA phonetic symbols 359

Google Exhibit 1008
Page 0013

B The EAGLET term database 367
B.1 Introduction 367
B.2 EAGLET termbank (abridged) 369

List of abbreviations 497

Index 503

CD–ROM disclaimer 521

List of Figures

1.1 An example of GToBI transcription, time-aligned with an F_0-track. 45
1.2 utt1 & utt2 76
1.3 utt3 & utt4 77
1.4 utt5 & utt6 78
1.5 utt7 79
1.6 utt8 & utt9 80
1.7 utt10 81
1.8 utt11 82
1.9 utt12 83

2.1 Model of human–computer interaction (from Schomaker et al. 1995a) 104
2.2 Modality-oriented classification of multimodal systems 115
2.3 Task-oriented taxonomy of multimodal applications 116
2.4 QuickDoc application – User gesture with speech input "This is Subdural Hematoma, confidence 90%" (from Waibel et al. 1997) 121
2.5 Multimodal Text Editor – User inserting the word "handwriting" by handwritten input (from Suhm 1997) 122
2.6 Taxonomy of system-level evaluation techniques, adopted from Balbo et al. (1993) 126
2.7 Late integration model (from Adjoudani et al. 1997) 132
2.8 Audio-visual time delay neural network (from Meier et al. 1997) 133
2.9 General flow chart of a talking head system (from Guiard-Marigny 1996) 134
2.10 Performance-based animation control (from Parke and Waters 1996) 136
2.11 Puppeteer animation control (from Parke and Waters 1996) . . . 136
2.12 Overview of an audio-visual speech system (from Adjoudani et al. 1997) 137
2.13 Multimodal Design Space (from Nigay and Coutaz 1993) 142
2.14 Fusion of two melting pots (from Nigay and Coutaz 1995) 144
2.15 Generalised input devices (from Schomaker et al. 1995a) 145
2.16 Application of the colour model to a sample input image. The face is marked in the input image. 158
2.17 Tracking of eyes, nostrils, and lips corners 162
2.18 Fascial tissue layers (from Parke and Waters 1996) 168
2.19 Parameterised facial model system (from Parke and Waters 1996) 169
2.20 Overview of a audio-visual rule-based system (from Beskow 1995) 170
2.21 Architecture of a system automatically generating an answer with the appropriate intonation and facial expression starting from a query (from Pelachaud and Prevost 1995) 174
2.22 Facial expression of imploration (from Pelachaud and Poggi 1998) 176
2.23 Facial expression accompanying the accented word 'amplifier' (from Pelachaud and Prevost 1995) 176
2.24 Handwriting recognition with explicit segmentation 182

2.25 Handwriting recognition with implicit segmentation 183
2.26 Architecture for a feature-based gesture recognition system . . . 186
2.27 Architecture for a 3D gesture recognition system 188

3.1 The diagnostic program EVAL 223

4.1 Extended semiotic triangle 244
4.2 Semiotic pyramid for multilingual termbases 244
4.3 Sign model 245
4.4 Example of a logical concept hierarchy 248
4.5 Example of an ontological hierarchy 248
4.6 Vauquois triangle for terminology 251
4.7 User interface of the TERMITE database 266
4.8 Interface of the EURODICAUTOM database 268
4.9 Structural overview of EAGLET 273
4.10 A basic corpus design taxonomy 275
4.11 A basic speech synthesis taxonomy 275
4.12 EAGLET Term Database interface 278
4.13 EAGLET administration interface 279

List of Tables

1.1 Sampson's subcategories for interjections 29
1.2 Some adverb subcategories from the London-Lund Corpus 30
1.3 Extended interjection POS categories. 30
1.4 ToBI Pitch Accents 42
1.5 Conversion between TSM and ToBI, according to Roach (1994) . 47
1.6 Tag definitions 74
1.7 Symbols for higher (non-terminal) constituents 75
1.8 English tagset, with intermediate tags 89
1.9 Italian DMI codes, with intermediate tags 95

2.1 Results from Survey of Multimodal Interfaces – Part I: Domain, Input/Output modalities, and Cooperation 199
2.2 Results from Survey of Multimodal Interfaces – Part II: Evaluation 201
2.3 Performance results of TDNN systems for speaker dependent (from Meier et al. 1997) 202
2.4 Results in word error (from Bregler and Konig 1994) 202
2.5 Prototype universal facial expressions of emotions and their corresponding FACS action units 202
2.6 List of AUs 203

3.1 Categorisation of some products and services into the categories described in Section 3.1.1. 205
3.2 The order of systems for different subjects. The numbers indicate the system number, time runs left to right. 212
3.3 Test order for a combination of two systems under two test conditions. The numbers indicate the order, e.g. the number 2 indicates that that (system,condition) is tested second for the test subject. 212
3.4 Examples of the use of a five-point scale 214
3.5 Recognition results, for three subjects. There is one condition where the enrolment speaker and the test speaker were not the same. The speaker who trained the exception vocabulary was always pp1. The test consisted of all 221 expressions defined in the word lists for ACT. The last column indicates the word error rate (WER), which does not include misses. 224
3.6 The recognition performance during the sorties, after cleaning up the utterances, and with a different recognition system. 226
3.7 Accuracy results in *C'T Magazin* test. (n.m. = not mentioned) . 231
3.8 Some results of the experiment, showing language dependence of the word accuracy, expressed in %. The last line shows the homophone error rate. 233
3.9 The penalty values for various errors 236

4.1 Relational database visualised as a table 262

Google Exhibit 1008
Page 0017

5.1 SLP related newsgroups 293
5.2 SLP related journals 295
5.3 WWW browsers 313
5.4 Web server 313

Editorial preface

The present volume is a companion to the *Handbook of Standards and Resources for Spoken Language Systems* (Gibbon et al. 1997), and addresses decision makers, developers and advanced students in the human language technologies who are looking for guidance in areas such as the evaluation of consumer off-the-shelf products, multimodal and audiovisual systems, terminology, and current resources for system development and evaluation.

The publication of the *Handbook* marked a milestone in the development of the spoken language technologies. Standards, in the sense of consensus about the best laboratory practice of the time, were formulated by leading experts and negotiated interactively with a large number of European laboratories working in the field. As with any handbook, even during the authoring phases it was clear that the *Handbook* could not keep pace with all developments in a rapidly expanding field, and the more modest aim of creating a standard common platform for best practice in the field was pursued. A mode of dissemination in which knowledge can be continually updated was considered; maintenance cost would, however, have been prohibitively high.

To the specialist, a number of gaps in the *Handbook* are evident, and full coverage of the field is currently an extremely elusive goal. The editors of the first *Handbook* selected a number of topics for further treatment, in consultation with project partners and other colleagues in the field. These topics formed the substance of work in the EAGLES Phase II Project (LE3-4244 10484/0), Work Packages 4, 5 and 6 (constituting the Spoken Language Working Group, SLWG), and were developed using the consensus negotiating procedures which had already been tried and tested during the development of the first EAGLES *Handbook*. The present volume presents the results of this work, which is indicated in the title *Handbook of Multimodal and Spoken Dialogue Systems: Resources, Terminology, and Product Evaluation*. Chairman of the SLWG was Roger Moore, coordinator and general editor SLWG was Dafydd Gibbon, and technical editor was Inge Mertins. Editorial workshops were held in Bielefeld, Germany, in January 1998 and in Leusden, Netherlands, in June 1998, attended by the SLWG technical authors and academic and industrial reviewers.

The Chapter *Representation and annotation of dialogue* was coordinated by Geoffrey Leech (Lancaster University, UK) and the main technical author was Martin Weisser (Lancaster University, UK) with substantive contributions from Andrew Wilson (Chemnitz University of Technology, Germany), and Martine Grice (Universität des Saarlandes, Saarbrücken, Germany). The Chapter was designed and edited in close consultation with Kerstin Fischer (Computer Science Department, Universität Hamburg, Germany), Susanne Jekat (Friedrich-Alexander-Universität, Erlangen-Nürnberg, Germany), Elisabeth Maier (SBC / IT Camp, Basel, Switzerland), Paul McKevitt (Aalborg University, Denmark), Jean Carletta (University of Edinburgh, UK), Joaquim Llisterri (Universitat Autònoma de Barcelona, Spain). Help from the following is also gratefully acknowledged: Anton Batliner, Niels Ole Bernsen, František Čermak, Alain Couillault, Paul Dalsgaard, Mika Enomoto, Ulrich Heid, Arne Johnsen, Magne Johnsen, Andreas Kellner, Gerry Knowles, Klaus Kohler, David Milward, Nor-

B The EAGLET term database

B.1 Introduction

For a comprehensive description of the principles and standards on which the EAGLET termbank is based, see Chapter 4. The present, slightly abridged printed version was automatically created from the EAGLET termbank database. In view of rapid development of the field and the degree of specialisation of many terms, the technical literature should be consulted for detailed definitions. In doubt, rather more general definitions have been preferred.
In the printed version the following data categories are retained:

1. Orthography
   A representation of the term in standard British English orthography.
2. Pronunciation
   The phonemic transcription of the term in both IPA and SAMPA notation is given. This is unusual for a termbank, but experience in the field shows a need among non-native experts for a pronunciation guide.
3. Part of Speech
   The structure of compounds is given in attribute–value notation. For example, the term 'text-to-speech system' is analysed as '[N: [N: text][PREP: to][N: speech][N: system]]'. The tags used here are to be read as:

   - N = noun
   - V = verb
   - AJ = adjective
   - AV = adverb
   - DET = determiner
   - PREP = preposition
   - C = conjunction
   - NU = numeral

9. Inflections
   As nearly all terms in EAGLET are nouns, this category basically indicates the plural form(s) of terms. The possible values are: -s ('badger' – 'badgers'), -es ('search' – 'searches'), none ('Bayesian decision theory'). For irregular forms and the '-ies' plural in words like 'frequencies' the plural form is given in full. When 'no plural' abstract nouns and generic names denote specific instances or types, they may take a plural in some contexts, e.g. 'three LPCs' in the sense of 'three LPC analyses'.
10. Domain
    'Domain' refers to the individual chapter of Gibbon et al. (1997) or of the present volume to which the term can be assigned or in which it is defined. Subject fields such as 'physical characterisation, 'corpora', 'lexicon', 'interactive dialogue systems' are specified. For example, for 'Hidden Markov Model' the value is 'Domain: language modelling'. Many terms, however, are difficult to place because they are very general, for instance 'orthographic transcription', a term that occurs in nearly all handbook chapters and, like many others, is not restricted to the domain of spoken language technology; in such cases the Domain field has been left empty.
11. Hyperonyms
    The data category 'hyperonym' corresponds to the classical *genera proximum*

of terminological and classical definition theory. A *hyperonym* is the verbal representation of the superordinate concept of a term in a taxonomy; it is the converse of 'hyponym'. For example, *morph* is a hyperonym of *bound morph* because '*A bound morph is a type of morph*' is an acceptable sentence. Similarly, *microphone* is a hyperonym of *unidirectional microphone*, and the latter is a hyperonym of *cardioid microphone*.

12. Hyponyms
    A *hyponym* is the verbal representation of the subordinate concept of the term in question; it is the converse of 'hyperonym'. Hyponyms of very general terms are omitted. Examples: A *bound morph* is a hyponym of *morph* because *A bound morph is a kind of morph* is an acceptable sentence.
13. Synonyms
    A synonym is a term that represents the same concept as the main entry term in a term entry. In EAGLET, no distinction is made between genuine synonyms and quasi-synonyms. Quasi-synonyms are terms that represent the same concept in the same language, but for which interchangeability is limited to some contexts and inapplicable in others. An example from the *Handbook*: *wolf* is a synonym of *skilled impostor*. Strictly speaking, abbreviations are synonyms for the terms they abbreviate, and are treated as such in the termbank.
14. Cohyponyms (cf. 'Antonym' in Chapter 4
    This data category covers antonyms, i.e. terms denoting various lexical opposites, without commitment to the particular kind of antonym. It includes complementaries, i.e. terms that "divide some conceptual domain into two mutually exclusive compartments" (Cruse 1986, p. 198). For example, in the spoken language technology domain, *recognition* and *synthesis* are cohyponyms (in fact, antonyms).
15. Definitions
    As in most standard general dictionaries, EAGLET not only contains analytical definitions, i.e. definitions which give a noun phrase which formulates the meaning of the term in question (Sager and L'Homme 1994), but also definitions that contain so-called 'nonessential' characteristics and information that would be classified as 'world knowledge'. In many cases the source of the definition is given.
16. Meronymic superordinates
    Terms that are superordinates in a PARTOF hierarchy. Example: *syllable* is a meronymic superordinate of *onset* because *The/An onset is part of a syllable* is an acceptable sentence.
17. Meronymic subordinates
    Terms that are subordinates in a PARTOF hierarchy. Example: *onset* is a meronym of *syllable*, because *An onset is a part of a syllable.* is an acceptable sentence.
18. Examples
    A term and its definition is exemplified.
    Example: 'un' and 'able' in 'unbearable' are affixes.

Google Exhibit 1008
Page 0021

B.2 EAGLET termbank (abridged)

2D gesture
/'tuɪ 'diɪ 'dʒestʃə/, /'tu: 'di: 'dZestS@/, [N: [AJ: 2D][N: gesture]], [plural: -s]. Domain: multimodal systems. Hyperonyms: gesture. Synonyms: graphic mark, gesture. Cohyponym: pointing, 3D gesture. Def.: 2D gestures refer to movements on a flat surface, for example marks drawn with a pen on a touch-sensitive display.

3D gesture
/'θriɪ 'diɪ 'dʒestʃə/, /'Tri: 'di: 'dZestS@/, [N: [AJ: 3D][N: gesture]], [plural: -s]. Domain: multimodal systems. Hyperonyms: gesture. Cohyponym: pointing, 2D gesture. Def.: 3D gestures refer to movements of fingers, hand, or head in three dimensional space.

abstract lemma
/'æbstrækt 'lemə/, /'{bstr{kt 'lem@/, [N: [AJ: abstract][N: lemma]], [plural: abstract lemmata]. Domain: lexicon. Cohyponym: lemma, lexical lemma. Def.: An abstract lemma is a lexical database access key which may have any convenient unique name or number (or be labelled by the spelling of the canonical inflected form); all properties have equal status, so that the abstract lemma is neutral with respect to different types of lexical access, through spelling, pronunciation, semantics, etc. (Gibbon et al. 1997, p. 200)

accent identification
/'æksənt aɪdentɪfɪ'keɪʃən/, /'{ks@nt aIdentIfI'keIS@n/, [N: [N: accent][N: identification]], [plural: -s]. Domain: speaker recognition. Hyperonyms: speaker classification task. Cohyponym: sex identification, age identification, mood identification, health state identification, speaker cluster identification. Def.: A task consisting in determining aspects of the social background of the speaker. (Gibbon et al. 1997, p. 409)

accent
/'æksənt/, /'{ks@nt/, [N: accent], [plural: -s]. Hyperonyms: 1. prosodic feature. Hyponyms: 1. syntactic accent, tonal accent. Synonyms: 1. pitch prominence. Def.: 1. The phonetic property which makes a particular word or syllable stand out in a stream of speech. (Crystal 1988, p. 2) 2. Regional, social or foreign pronunciation.

acceptance
/ək'septəns/, /@k'sept@ns/, [N: acceptance], [plural: -s]. Domain: speaker recognition. Hyperonyms: decision outcome (of a speaker recognition system). Hyponyms: false acceptance. Cohyponym: rejection. Def.: 1. Decision outcome which consists in responding positively to a task such as a speaker (or speaker class) verification task 2. The degree to which customers are willing to use a system or service.

accuracy
/'ækjʊrəsi/, /'{kjUr@si/, [N: accuracy], [plural: none]. Domain: speaker recognition. Hyperonyms: performance measure. Hyponyms: recognition accuracy. Synonyms: precision. Cohyponym: recall; error rate. Def.: A measure of the performance of a system such as an automatic speech recognition (ASR) system, defined as (N - S - D - I)/N, where N: number of basic units (usually words) in a test, S: number of substitution errors, D: number of deletion errors, I: number of insertion errors.

acoustic interface
/ə'kuɪstɪk 'ɪntəfeɪs/, /@'ku:stIk 'Int@feIs/, [N: [AJ: acoustic][N: interface]], [plural: -s] Domain: speech synthesis, speech recognition. Cohyponym: linguistic interface. Meronym. sup.: text-to-speech system. Def.: 1. The acoustic interface of a speech synthesiser transduces the output of the linguistic interface (lexical representation, abstract phonological code) to an audible waveform. 2. The acoustic interface of a speech recogniser converts the acoustic input signal into a set of word or sentence hypotheses.

Google Exhibit 1008
Page 0022

acoustic measure
/əˈkuːstɪk ˈmeʒə/, /@'ku:stIk 'meZ@/, [N:[AJ: acoustic][N: measure]], [plural: -s]. Hyponyms: amplitude, intensity, fundamental frequency, F0. Def.: An acoustic measure quantifies properties of a system on the basis of properties of the speech signal it processes.

acoustic output device
/əˈkuːstɪk ˈaʊtpʊt dɪˈvaɪs/, /@'ku:stIk 'aUtpUt dI'vaIs/, [N: [AJ: acoustic][N: output][N: device]], [plural: -s]. Domain: multimodal systems. Hyperonyms: output device. Cohyponym: visual output device, haptic output device. Def.: Much research has been conducted into producing good quality synthetic speech, and there are several commercial products. Non-speech sounds include beep sounds, auditory icons or earcons, and auditory display (visualisation of data through sound parameters). Virtual reality systems or headphones can simulate the spatial relations of sounds.

acoustic phonetics
/əˈkuːstɪk fəˈnetɪks/, /@'ku:stIk f@'netIks/, [N: [AJ: acoustic][N: phonetics]], [plural: none]. Hyperonyms: phonetics. Cohyponym: articulatory phonetics, auditory phonetics. Meronym. sup.: phonetics. Def.: Acoustic phonetics is the study of the physical properties of speech sound, as transmitted between mouth and ear. (Crystal 1988)

acoustic-phonetic model
/əˈkuːstɪk fəˈnetɪk ˈmɒdəl/, /@'ku:stIk f@'netIk 'mQd@l/, [N: [AJ: acoustic][AJ: phonetic][N: model]], [plural: -s]. Domain: language modelling. Hyperonyms: knowledge source in a speech recognition system; model. Cohyponym: language model. Meronym. sup.: automatic speech recognition system. Def.: The acoustic-phonetic model is the conditional probability of observing the acoustic vectors when the speaker utters the words. Like the language model probabilities, these probabilities are estimated during the training phase of the recognition system. (Gibbon et al. 1997, p. 239)

acquainted impostor
/əˈkweɪntɪd ɪmˈpɒstə/, /@'kweIntId Im'pQst@/, [N: [AJ: acquainted][N: impostor]], [plural: -s]. Domain: speaker recognition. Hyperonyms: intentional impostor. Cohyponym: unacquainted impostor. Def.: An acquainted impostor qualifies as an intentional impostor who has some knowledge of the voice of the authorised speaker. (Gibbon et al. 1997, p. 422)

ACT
1. /ˈeɪ ˈsiː ˈtiː/ 2. /ˈækt/, 1. /'eI 'si: 'ti:/ 2. /'{kt/, [N: ACT], [plural: -s]. Domain: consumer off-the-shelf products. Hyperonyms: tool. Synonyms: Advanced Crew Terminal. Def.: The ACT is a collection of tools that can help an astronaut in his daily work, providing electronic time schedules, procedure checking, experiment control and data acquisition. It was inplemented in a Microsoft Windows operating environment as a collection of application programs.

Action Unit
/ˈækʃən ˈjuːnɪt/, /'{kS@n 'ju:nIt/, [N: [N: Action][N: Unit]], [plural: -s]. Domain: multimodal systems. Hyperonyms: unit. Synonyms: AU. Def.: Basic unit used in FACS. An AU corresponds to the action of a muscle or a group of related muscles. Each AU describes the direct effect of muscle contraction as well as any secondary effects due to movement propagation, wrinkles or bulges. A facial expression is the combination of AUs. Most of the AUs combine additively. But they may also be subject to rules of dominance (an AU disappears for the benefit of another AU), substitution (an AU is eliminated when others produce the same effect), alteration (AUs cannot combine).

Google Exhibit 1008
Page 0023

active contour
/'æktɪv 'kɒntʊə/, /'{ktIv 'kQntU@/, [N: [AJ: active][N: contour]], [plural: -s]. Domain: multimodal systems. Synonyms: snake. Def.: Deformable contour defined by a set of nodes connected by springs. Active contours are first located on the face. Contours are tracked by applying an image force field that is computed from the gradient of the intensity image. Muscle contraction is estimated from contour deformations. The import of visual information to recognise audio signals is around 7 percent.

active vocabulary size
/'æktɪv və'kæbjʊləri 'saɪz/, /'{ktIv v@'k{bjUl@ri 'saIz/, [N: [AJ: active][N: vocabulary][N: size]], [plural: -s]. Domain: speech recognition, consumer off-the-shelf products. Hyperonyms: coverage, vocabulary size. Cohyponym: passive vocabulary size; exception vocabulary size, extension vocabulary size, user vocabulary size. Def.: A maximum number of words a speech recognition system can recognise at any given moment.

active vocabulary
/'æktɪv və'kæbjʊləri/, /'{ktIv v@'k{bjUl@ri/, [N: [AJ: active][N: vocabulary]], [plural: y/-ies]. Domain: system design. Hyperonyms: lexicon, vocabulary. Cohyponym: backup vocabulary. Def.: The vocabulary that is actively considered by a recogniser at a given instant.

activity type
/æk'tɪvɪti 'taɪp/, /{k'tIvIti 'taIp/, [N: [N: activity][N: type]], [plural: -s]. Domain: interactive dialogue systems. Hyperonyms: communicative activity. Def.: A type of communicative activity with its own conventions of dialogue. E.g. negotiation, instruction, problem solving.

actual word
/'ækʃʊəl 'wɜːd/, /'{kSU@l 'w3:d/, [N: [AJ: actual][N: word]], [plural: -s]. Domain: lexicon. Hyperonyms: word, lexical item. Synonyms: lexicalised word. Cohyponym: potential word, neologism, nonce word. Def.: Word contained in a finite lexicon (Gibbon et al. 1997, p. 195)

adaptive dialogue strategy
/ə'dæptɪv 'daɪəlɒg 'strætədʒi/ , /@'d{ptIv 'daI@lQg 'str{t@dZi/ , [N: [AJ: adaptive][N: dialogue][N: strategy]], [plural: y/-ies]. Domain: interactive dialogue systems. Hyperonyms: dialogue strategy. Cohyponym: constitutive dialogue strategy, cooperative dialogue strategy, deterministic dialogue strategy. Def.: An adaptive dialogue strategy takes into account a dynamic user model by learning the user's communicative strategies and adjusting to them as each dialogue proceeds. (Gibbon et al. 1997, p. 598)

adaptive language model
/ə'dæptɪv 'læŋgwɪdʒ 'mɒdəl/, /@'d{ptIv 'l{NgwIdZ 'mQd@l/, [N: [AJ: adaptive][N: language][N: model]], [plural: -s]. Domain: language modelling. Hyperonyms: language model. Cohyponym: non-adaptive language model. Def.: Adaptive language models adapt their probabilities to the most recent history, say the last 100 to 1000 predecessor words. (Gibbon et al. 1997, p. 257)

adequacy evaluation
/'ædɪkwəsi ɪvæljʊ'eɪʃən/, /'{dIkw@si Iv{ljU'eIS@n/, [N: [N: adequacy][N: evaluation]], [plural: -s]. Domain: multimodal systems. Hyperonyms: evaluation. Cohyponym: diagnostic evaluation, performance evaluation, comparative assessment, benchmarking assessment. Def.: Adequacy evaluations determine the fitness of a system for a purpose: does it meet the requirements, how well, and at what cost? The requirements are mainly determined by user needs. Therefore, user needs have to be identified, which may require considerable effort in itself. Consumer reports are a typical example of adequacy evaluation.

Google Exhibit 1008
Page 0024

complementarity
/kɒmplɪmən'tærɪti/, /kQmplIm@n't{rIti/, [N: complementarity], [plural: none]. Domain: multimodal systems. Hyperonyms: cooperation type. Cohyponym: redundancy, equivalence, specialisation, concurrency, transfer. Def.: Different chunks of information belonging to the same command are transmitted over more than one modality. E.g. Saying "put-that-there", while pointing at an object, and then at a location..

component evaluation
/kəm'pəʊnənt ɪvæljʊ'eɪʃən/, /k@m'p@Un@nt Iv{ljU'eIS@n/, [N: [N: component][N: evaluation]], [plural: -s]. Domain: multimodal systems. Hyperonyms: evaluation. Def.: Evaluation of the components of a system such as a multimodal system. Evaluation methodologies that are accepted in the various subfields can be reused, including evaluation of speech recognition, handwriting recognition, and gesture recognition, as well as the evaluation of talking heads. In addition, the quality of the integration of the components in a multimodal system may have to be evaluated, for example, the accuracy of automatically assigning multimodal input to the appropriate (specialised) recognisers.

composite word
/'kɒmpɒsɪt 'wɜːd/ , /'kQmpQsIt 'w3:d/ , [N:[AJ: composite][N: word]], [plural: -s]. Domain: word formation. Hyperonyms: word. Hyponyms: compound word, derived word. Cohyponym: simplex word. Def.: A composite word is either a compound word or a derived word.

compound
/'kɒmpaʊnd/, /'kQmpaUnd/, [N: compound], [plural: -s]. Domain: lexicon. Hyperonyms: word. Synonyms: composite word. Cohyponym: simplex word. Def.: A compound is a word morphologically concatenated with a word or stem and thus contains at least two stems.

compounding
/'kɒmpaʊndɪŋ/, /'kQmpaUndIN/, [N: compounding], [plural: none]. Domain: lexicon. Hyperonyms: morphological operation. Synonyms: composition. Cohyponym: derivation. Meronym. sup.: word formation. Def.: Compounding deals with the construction of words by concatenating words or stems. (Gibbon et al. 1997, p. 214) E.g. wind + mill = windmill.

comprehension test
/kɒmprɪ'henʃən 'test/, /kQmprI'henS@n 'test/, [N: [N: comprehension][N: test]], [plural: -s]. Domain: speech synthesis. Hyperonyms: testing procedure. Hyponyms: off-line comprehension test, on-line comprehension test. Cohyponym: identification test. Def.: Procedure testing a listener's understanding of a speech stimulus at the sentence or text level (often by asking the listener to answer content questions).

compression scheme
/kəm'preʃən 'skiːm/, /kQm'preS@n 'ski:m/, [N: [N: compression][N: scheme]], [plural: -s]. Hyponyms: MPEG-1, MPEG-2, MPEG-3, MPEG-4, Huffman compression. Def.: A compression scheme is an algorithm for reducing the size of a data packet, for example by searching for and indexing repeated identical sequences.

concatenation technique
/kənkætə'neɪʃən tek'niːk/, /k@nk{t@'neIS@n tek'ni:k/, [N: [N: concatenation][N: technique]], [plural: -s]. Domain: speech synthesis, consumer off-the-shelf products. Hyperonyms: speech synthesis technique. Cohyponym: production model, playback technique. Def.: The technique of concatenating and smoothing pre-defined segments of speech for speech synthesis. By playing back sub-word units of pre-recorded speech contiguously, whole words and phrases can be synthesised. Mostly, the units chosen are diphones, i.e. the period of the last half of the previous phone up to the first half of the next phone. Usually, the voice quality of these systems is high. By using techniques such as PSOLA the pitch of the pre-recorded waveforms can be changed and thus controlled intonation and stress is possible. A genuine change of voice characteristics is not possible. The vocabulary is limited by pronunciation rules.

Google Exhibit 1008
Page 0025

concatenative synthesis
/kɒnˈkætɪnətɪv ˈsɪnθəsɪs/, /kQn'k{tIn@tIv 'sInT@sIs/, [N: [AJ: concatenative][N: synthesis]], [plural: concatenative syntheses]. Domain: speech synthesis. Hyperonyms: speech synthesis. Cohyponym: parametric synthesis. Def.: Speech synthesis where samples of predefined segments of speech are concatenated, usually with some smoothing applied so that the boundaries are less audible.

concept-to-speech system
/ˈkɒnsept ˈtə ˈspiːtʃ ˈsɪstəm/, /'kQnsept 't@ 'spi:tS 'sIst@m/, [N: [N: concept][PREP: to][N: speech][N: system]], [plural: -s]. Domain: speech synthesis. Hyperonyms: speech output system. Synonyms: CTS, meaning to speech, MTS. Cohyponym: text-to-speech system, TTS. Def.: A concept-to-speech system is a speech output system that converts some abstract representation of a communicative intention to speech rather than a text representation.

conceptual appropriateness
/kənˈseptjʊəl əˈprəʊprɪətnɪs/, /k@n'septjU@l @'pr@UprI@tnIs/, [N: [AJ: conceptual][N: appropriateness]], [plural: none]. Domain: interactive dialogue systems. Hyperonyms: measure. Def.: Conceptual appropriateness is a measure of the appropriateness of a system utterance in its immediate dialogue context. This is a five-valued measure, with values drawn from one set: - TF (total failure) - AP (appropriate) - IA (inappropriate) - AI (appropriate / inappropriate)- IC (incomprehensible). (Gibbon et al. 1997, p. 606/607)

concurrency
/kənˈkʌrənsi/, /k@n'kVr@nsi/, [N: concurrency], [plural: none]. Domain: multimodal systems. Hyperonyms: cooperation type. Cohyponym: complementarity, redundancy, equivalence, specialisation, transfer. Def.: Independent chunks of information are transmitted using different modalities and overlap in time. Concurrency means parallel use of different modalities to initiate different actions. E.g. Talking over speaker phone while editing a document..

condenser microphone
/kɒnˈdensə ˈmaɪkrəfəʊn/, /kQn'dens@ 'maIkr@f@Un/, [N: [N: condenser][N: microphone]], [plural: -s]. Domain: physical characterisation. Hyperonyms: microphone. Cohyponym: dynamic microphone. Def.: Condenser microphones basically consist of a capacitor, one of the electrodes of which is formed by a conductive membrane. This membrane is exposed to the incident sound and, when moved back and forth by the sound pressure, slightly changes the capacitance of the capacitor. When the load on the capacitor is kept constant the capacitance changes will, for the voltage across the electrodes, follow the movements of the membrane as long as the voltage changes are small compared to the total voltage across the electrodes. Since the membrane can be manufactured from very thin plastic film material with a conductive layer of vapourised gold or aluminium, it will follow the sound pressure quite exactly and the signal produced by the microphone will be a rather precise reproduction of the original course of the sound pressure. For high-quality studio recordings most microphones used are condenser microphones. (Gibbon et al. 1997, p. 302/303)

confidence interval
/ˈkɒnfɪdəns ˈɪntəvəl/, /'kQnfId@ns 'Int@v@l/, [N: [N: confidence][N: interval]], [plural: -s]. Domain: assessment methodologies. Hyperonyms: statistical measure. Def.: A statistical measure specifying the proportion of independent estimates of some population value that are likely to fall within this interval.

confidence measure
/ˈkɒnfɪdəns ˈmeʒə/, /'kQnfId@ns 'meZ@/, [N: [N: confidence][N: measure]], [plural: -s]. Hyperonyms: value, measure. Def.: A value computed, for instance, by an automatic speech recogniser indicating the degree of belief that the word, phrase or sentence to which the confidence measure was assigned has been recognised correctly.

confidence
/ˈkɒnfɪdəns/, /'kQnfId@ns/, [N: confidence], [plural: none]. Domain: speaker recognition. Def.: The degree of belief in the correctness of a decision made by a system.

Google Exhibit 1008
Page 0026

identity assignment
/aɪˈdentɪti əˈsaɪnmənt/, /aI'dentIti @'saInm@nt/, [N: [N: identity][N: assignment]] , [plural: -s] . Domain: speaker recognition. Hyperonyms: decision outcome. Def.: Decision outcome which consists in attributing an identity to an applicant speaker, in the context of speaker identification. For speaker classification, the term class assignment should be used instead.

idiolect
/ˈɪdɪəlekt/, /'IdI@lekt/, [N: idiolect], [plural: -s]. Hyperonyms: language variety. Meronym. sup.: natural language. Def.: Idiolect is a term used in linguistics to refer to the linguistic system of an individual speaker - his personal dialect. (Crystal 1988) An idiolect is heterogeneous, i.e. a speaker switches between different codes (code-switching) and shifts between different styles of speaking (style-shifting).

idiom
/ˈɪdɪəm/, /'IdI@m/, [N: idiom], [plural: -s]. Domain: lexicon. Hyperonyms: lexical unit. Hyponyms: pragmatic idiom, phrasal idiom. Def.: 1. Lexical unit larger than the word whose meaning is not derivable from the meanings of the composite words, which has restrictions on alternative word orders and possible component words, and may not be strictly grammatical. 2. A variety of a language (e.g. a dialect). E.g. 'Come to think of it, ...' in the meaning 'I just thought of another relevant point, namely ...'.

impostor
/ɪmˈpɒstə/, /Im'pQst@/, [N: impostor], [plural: -s] . Domain: speaker recognition. Hyperonyms: applicant speaker. Hyponyms: intentional impostor, well-intentioned impostor. Synonyms: impersonator, usurper, usurpator.(Both terms are very rarely used.), non-registered speaker (in identification) . Cohyponym: registered speaker. Def.: In speaker identification: an applicant speaker who does not belong to the set of registered speakers. In speaker verification: a speaker whose real identity is different from his claimed identity. (Gibbon et al. 1997, p. 413)

infix
/ˈɪnfɪks/, /'InfIks/, [N: infix], [plural: -es]. Domain: lexicon. Hyperonyms: affix. Cohyponym: prefix, suffix, interfix, superfix, circumfix. Meronym. sup.: word. Def.: Morph that is inserted into a stem. E.g. Latin 'iungere' vs. 'iugum'.

inflection
/ɪnˈflekʃən/, /In'flekS@n/, [N: inflection], [pural: -s]. Domain: lexicon. Hyperonyms: morphological operation. Cohyponym: word formation. Meronym. sup.: morphology. Def.: Inflection is the branch of morphology which deals with markers of the relation of words to their contexts within sentences, on the basis of agreement (congruence), e.g. between subject and verb. (cf. also Gibbon et al. 1997, p. 214) E.g. The fish tastes good. The final 's' in 'tastes' indicates that the subject noun is '3rd Person Singular', while in The fish taste good, the lack of a final 's' indicates that the subject noun, which is 3rd person, is plural.

inflectional affixation
/ɪnˈflekʃənəl æfɪkˈseɪʃən/, /In'flekS@n@l {fIk'seIS@n/, [N; [A.J: inflectional][N: affixation]], [plural: -s]. Domain: lexicon. Hyperonyms: affixation. Hyponyms: inflectional prefixation, inflectional suffixation. Cohyponym: derivational affixation. Def.: Morphological concatenation of a stem with a full set of inflectional affixes. (Gibbon et al. 1997, p. 215) E.g. English 'algorithm' + 's' = 'algorithms'; German 'ge' + 'segn' + 'et' + 'en' 'blessed' (plural participle or adjective). (Gibbon et al. 1997, p. 215).

Google Exhibit 1008
Page 0027

input modality
/ˈɪnpʊt məˈdælɪti/, /'InpUt m@'d{lIti/, [N: [N: input][N: modality]], [plural: y/-ies]. Domain: multimodal systems. Hyperonyms: modality. Hyponyms: non-speech input modality. Cohyponym: output modality. Def.: A human sensory input channel used in communication. In human-machine communication, when input from a system the input modalities are generally acoustic (loudspeaker, headphones) or visual (monitor screen), but may be haptic or tactile (e.g. Braille tablets). The olfactory (smell) and gustatory (taste) senses are additionally used in human-human communication.

insertion
/ɪnˈsɜːʃən/, /In's3:S@n/, [N: insertion], [plural: -s]. Domain: speech recognition, language modelling. Synonyms: false alarm. Cohyponym: deletion, substitution. Def.: A response to a word that was not in the utterance presented to a speech recognition system (or false alarm).

Integrated Services Digital Network
/ˈɪntɪgreɪtɪd ˈsɜːvɪsɪz ˈdɪdʒɪtəl ˈnetwɜːk/, /'IntIgreItId 's3:vIsIz 'dIdZIt@l 'netw3:k/, [N: [AJ: Integrated][N: Services][AJ: Digital][N: Network]], [plural: none]. Hyperonyms: digital telephony. Synonyms: ISDN. Cohyponym: Global System for Mobile Communication, GSM; analogue telephony. Def.: A world-wide standard for digital telephony in fixed networks.

intensity
/ɪnˈtensɪti/, /In'tensIti/, [N: intensity], [plural: y/-ies]. Domain: physical characterisation. Hyperonyms: acoustic measure. Cohyponym: amplitude, fundamental frequency, F0. Def.: Intensity is power per unit area, or the way power is distributed in a space. Power itself is a measure of the rate at which energy is being expended ... Now it can be shown that intensity is proportional to the square of pressure. (Crystal 1988, p. 223)

intentional impostor
/ɪnˈtenʃənəl ɪmˈpɒstə/, /In'tenS@n@l Im'pQst@/, [N: [AJ: intentional][N: impostor]], [plural: -s]. Domain: speaker recognition. Hyperonyms: impostor. Hyponyms: acquainted impostor, unacquainted impostor. Cohyponym: well-intentioned impostor. Def.: An intentional impostor has the clear goal of being identified or verified though he is not registered (violation), or to be identified as somebody else (usurpation). (Gibbon et al. 1997, p. 422)

interaction
/ɪntərˈækʃən/, /Int@r'{kS@n/, [N: interaction], [plural: -s]. Hyperonyms: communication. Def.: Communication of information between two agents, in which (except for the special case of the initial turn) an agent's contribution at any given point can be construed as a response to the previous turn or turns. A signal or a stimulus coming from one agent provokes a change in the internal state(s) of, and a response(s) from, the other agent. Stimuli are understood in a broad sense, including multimodal stimuli (in which different media may be used). They may, for example, consist of a physical action (moving a pointer) or a linguistic act (uttering a sentence) or the coordination of both.

interactive dialogue system
/ɪntərˈæktɪv ˈdaɪəlɒg ˈsɪstəm/ , /Int@r'{ktIv 'daI@lQg 'sIst@m/ , [N: [AJ: interactive][N: dialogue][N: system]], [plural: -s] . Domain: interactive dialogue systems. Hyperonyms: dialogue system. Synonyms: Interactive Voice Response; spoken language dialogue system. Cohyponym: command system. Def.: A computer system capable of engaging in turn-by-turn communication with a human user. In the general case, communication between the person and the system could use any communication mode or medium (or several simultaneously). In this chapter, however, the term is usually used more restrictively to apply to systems whose primary mode of communication is spoken natural language. This interactive communication is based on the integration of a set of modules, each of which handles a complex task. The modules are linked to each other and interactions are controlled by a kernel module which has the overall task of managing the dialogue. (Gibbon et al. 1997, page 567 / 568)

Google Exhibit 1008
Page 0028

iterative design
/ˈɪtərətɪv dɪˈzaɪn/, /'It@r@tIv dI'zaIn/, [N: [AJ: iterative][N: design]], [plural: none]. Hyperonyms: experimental technique. Synonyms: rapid prototyping. Cohyponym: benchmark evaluation, user study, simulation study. Def.: Iterative design has been widely adopted in the field of human-computer interaction, especially for product development. It is suitable for the development of multimodal applications, since many detail implementation issues can be explored rather quickly. The iterative design cycle includes (re)design of the application, implementation, and (informal) user testing. Iterative design is highly desirable from the HCI point of view but is difficult to reconcile with the pipeline or cascaded process organisation in software development which is currently still predominant, for reasons of cost control mainly.

IVR
/ˈaɪ ˈviː ˈɑː/, /'aI 'vi: 'A:/, [N: IVR], [plural: -s] . Domain: interactive dialogue systems, system design. Hyperonyms: dialogue system. Synonyms: interactive voice response, interactive dialogue. Def.: Interactive Voice Response (IVR) is what the commercial world calls interactive dialogue. As such, its scope encompasses certain kinds of simple interaction which research scientists do not normally think of as dialogues. For example, a telephone caller calling a weather forecasting Audiotex service might be asked to say one of the words "today", "tomorrow" or "weekend". In the basis of what the system recognises, a canned weather forecast will be played. This is an example of IVR which is also widely known as Voice Response (VR), and a system which supports VR is usually known as a Voice Response Unit (VRU).

jacknife method
/ˈdʒæknaɪf ˈmeθəd/, /'dZ{knaIf 'meT@d/, [N: [N: jacknife][N: method]], [plural: -s]. Hyperonyms: testing method. Def.: A method of testing a speech recognition system, for example, in which part of the data is kept aside for testing and the rest used for training, and the part kept aside is rotated until all the data has been used for testing.

Java
/ˈdʒɑɪvə/, /'dZA:v@/, [N: Java], [plural: none]. Hyperonyms: object-oriented programming language. Def.: An object-oriented programming language developed by Javasoft of SUN Microsystems. It has become the de facto standard programming language for applets, i.e. programs which are distributed over the WWW to run inside a WWW browser. Java features a large class library including classes for graphical display, and audio data access. The Java Speech API specification supports voice command recognition, dictation, and text-to-speech synthesis.

jitter
/ˈdʒɪtə/, /'dZIt@/, [N: jitter], [plural: -s]. Domain: physical characterisation. Hyperonyms: physical characteristic; frequency modulation. Cohyponym: wow. Def.: Jitter is a form of frequency modulation of a signal by noise. The term is used as a measure of the average perturbation of a speaker's fundamental frequency and of its magnitude.

judgment testing
/ˈdʒʌdʒmənt ˈtestɪŋ/ , /'dZVdZm@nt 'testIN/ , [N: [N: judgment][N: testing]] , [plural: none] . Domain: speech synthesis. Hyperonyms: testing procedure. Synonyms: opinion testing. Cohyponym: functional testing. Def.: Procedure whereby a group of listeners is asked to judge the performance of a speech output system along a number of rating scales.

J_ToBI
/ˈdʒeɪ ˈtəʊbi/, /'dZeI 't@Ubi/, [N: [N: J_ToBI]], [plural: none]. Domain: corpora. Hyperonyms: ToBI . Cohyponym: E_ToBI, GlaToBI, G_ToBI. Def.: J_ToBI is a variant of ToBI developed for the transcription of Standard (Tokyo) Japanese.

keyword detection
/ˈkiːwɜːd dɪˈtekʃən/, /'ki:w3:d dI'tekS@n/, [N: [N: keyword][N: detection]], [plural: -s]. Synonyms: word spotting. Def.: The detection of a given word or word sequence in a stream of speech or text.

Google Exhibit 1008
Page 0029

knowledge-based speech recognition system
/nɒlɪdʒ'beɪst 'spiːtʃ rekəg'nɪʃən 'sɪstəm/, **/nQlIdZ'beIst 'spi:tS rek@g'nIS@n 'sIst@m/**, [N: [N: knowledge][AJ: based][N: speech][N: recognition][N: system]], [plural: -s]. Domain: speech recognition, corpora. Hyperonyms: speech recognition system. Cohyponym: stochastic speech recognition system. Def.: A knowledge-based speech recognition system specifies explicit acoustic-phonetic rules that are robust enough to allow recognition of linguistically meaningful units and that ignore irrelevant variation in these units. (Gibbon et al. 1997, p. 94)

labiodental consonant
/leɪbɪəʊ'dentəl 'kɒnsənənt/, **/leIbI@U'dent@l 'kQns@n@nt/**, [N: [AJ: labiodental][N: consonant]], [plural: -s]. Hyperonyms: consonant. Cohyponym: bilabial consonant, dental consonant, alveolar consonant, postalveolar consonant, retroflex consonant, palatal consonant, velar consonant, uvular consonant, pharyngeal consonant, glottal consonant. Def.: A labiodental consonant is a consonant classified on the basis of its place of articulation: a sound in which one lip is actively in contact with the teeth. (Crystal 1988, p. 172)

laboratory room
/lə'bɒrətri 'rʊm/, **/l@'bQr@tri 'rUm/**, [N: [N: laboratory][N: room]], [plural: -s]. Domain: physical characterisation. Hyperonyms: recording room. Cohyponym: soundproof booth, recording studio, anechoic chamber. Def.: Speech recordings in typical laboratory environments are sometimes made in a kind of workbench situation when no special recording facility is available. Recordings made in a laboratory environment are often used to test speech recognition systems, as lab speech recordings seem to reflect best natural speech recognition situations, without requiring too much effort concerning the recording setup. (Gibbon et al. 1997, p. 309)

laboratory testing
/lə'bɒrətri 'testɪŋ/, **/l@'bQr@tri 'testIN/**, [N: [N: laboratory][N: testing]], [plural: none] . Domain: speech synthesis. Hyperonyms: testing procedure. Cohyponym: field testing. Def.: Speech output test procedure entirely run in a laboratory, either abstracting from in vivo complications or trying to simulate real-life situations.

lamb
/'læm/, **/'l{m/**, [N: lamb], [plural: -s]. Domain: speaker recognition. Hyperonyms: speaker. Synonyms: vulnerable speaker. Cohyponym: resistant speaker. Def.: A speaker with a high mistrust rate. (Gibbon et al. 1997, p. 433)

language disorder
/'læŋgwɪdʒ dɪs'ɔːdə/, **/'l{NgwIdZ dIs'O:d@/**, [N: [N: language][N: disorder]], [plural: -s]. Domain: corpora. Hyperonyms: speech disorder. Synonyms: aphasia. Cohyponym: articulation disorder, resonance disorder, voice disorder, rhythm disorder. Def.: A language disorder is one which does not affect only the production of the speech message, but rather its content. (Gibbon et al. 1997, p. 115)

Google Exhibit 1008
Page 0030

language model
/'læŋgwɪdʒ 'mɒdəl/, /'l{NgwIdZ 'mQd@l/, [N: [N: language][N: model]], [plural: -s] . Domain: language modelling. Hyperonyms: model; knowledge source (in a speech recognition system); formal representation (of the structure of a natural language). Hyponyms: uniform language model; finite state language model; stochastic language model, grammar based language model; bigram language model, trigram language model, n-gram language model. Cohyponym: acoustic-phonetic model. Meronym. sup.: automatic speech recognition system. Def.: A formal representation of the structure (usually the sentence structure) of a natural language. In most speech recognisers language models consist of an n-gram model, i.e. probabilities of the occurrence of words, words pairs, sequences of three words, etc. In NLP language models tend to consist of formal grammars, which are sometimes enriched by statistical information. A language model in speech recognition is used to improve the recognition accuracy. Its task is to capture the redundancy inherent to the word sequences to be recognised. This redundancy may result from both the task specific constraints and general linguistic constraints. (Gibbon et al. 1997, p. 845)

language variety
/'læŋgwɪdʒ və'raɪəti/, /'l{NgwIdZ v@'raI@ti/, [N:[N: language][N: variety]], [plural: y/-ies]. Domain: . Hyponyms: dialect, idiolect, register, sociolect, sublanguage, vernacular. Def.: A language variety is one of a family of reasonably homogeneous situation-dependent codes used by a group of speakers and recognised by them to belong to a single language. Language varieties can be defined in terms of their location in a variety space with three main dimensions: regional variation (dialect), social variation (sociolect), functional variation (register, style, sublanguage, technical language). In general, systems are designed for speaker-dependent or speaker-independent performance within one language variety only.

language
/'læŋgwɪdʒ/, /'l{NgwIdZ/, [N: language], [plural: -s]. Domain: . Hyponyms: formal language, natural language. Def.: 1. Language is the faculty underlying the human ability to communicate by means of intricate sound patterns. 2. A language is a system of signs defined by a syntax (a vocabulary and a set of combinatory rules), a semantics (a set of rules for assigning meanings to vocabulary elements and their combinations), and a pragmatics (a set of conventions for using the vocabulary elements and their combinations in specific situations).

laryngograph
/lə'rɪŋgəgrɑːf/, /l@'rINg@grA:f/, [N: laryngograph], [plural: -s]. Domain: corpora, physical characterisation. Hyperonyms: device. Synonyms: electroglottograph. Def.: A laryngograph is a device for measuring the high frequency impedance across the glottis during phonation; the resulting signal (known as Lx) is closely related to the rate of articulation (in the articulatory domain), the fundamental frequency (in the acoustic domain), and the impression of pitch (in the auditory domain). In general, Lx is a closer and more noise-free approximation to phonetially relevant patterns than acoustically measured measurement of fundamental frequency. Laryngograph is a proprietary term.

larynx
/'lærɪŋks/, /'l{rINks/, [N: larynx], Domain: corpora, physical characterisation. Hyperonyms: articulator. Def.: Larynx is the part of the windpipe containing the vocal cords. (Crystal 1988, p. 174)

late integration
/'leɪt ɪntɪ'greɪʃən/, /'leIt IntI'greIS@n/, [N:L [AJ: late][N: integration]], [plural: -s]. Domain: multimodal systems. Cohyponym: early integration. Def.: Integration of audio and visual information in HMMs where a first decision based on the separate signals is taken and then the final decision based on the combination of both results is made.

Google Exhibit 1008
Page 0031

lateral
/'lætərəl/, /'l{t@r@l/, [N: lateral], [plural: -s]. Hyperonyms: consonant. Def.: Lateral is a term used in phonetic classification of consonant sounds on the basis of their manner of articulation: it refers to any sound where the air escapes around one or both sides of a closure made in the mouth. E.g. Example in English: [l].

learning effect
/'lɜːnɪŋ ɪ'fekt/, /'l3:nIN I'fekt/, [N: [N: learning][N: effect]], [plural: -s]. Domain: speech recognition, consumer off-the-shelf products. Def.: The effect that a first test for a naive subject is always harder than later tests, because the subject learns how to deal with the system, what kind of events to expect, etc.

leaving-one-out
/'liːvɪŋ 'wʌn 'aʊt/, /'li:vIN 'wVn 'aUt/, [N: [V: leaving][DET: one][PREP: out]], [plural: none]. Domain: language modelling. Hyperonyms: cross-validation. Def.: Leaving-one-out is a special kind of cross-validation where no additional test set is needed. Instead, it is generated from the training observations by leaving out one observation at a time.

left-to-right coarticulation
/'left tə 'raɪt kəʊɑːtɪkjʊ'leɪʃən/, /'left t@ 'raIt k@UA:tIkjU'leIS@n/, [N: [AJ: left][PREP: to][AJ: right][N: coarticulation]], [plural: none]. Hyperonyms: coarticulation. Synonyms: perseverative coarticulation, forward coarticulation. Cohyponym: backward coarticulation, anticipatory coarticulation, right-to-left coarticulation. Def.: In the string ...AB..., sound A influences sound B (or beyond). L>R coarticulation is thought to be largely due to lag in articulatory movement, induced by inertia. (Clark & Yallop, p. 87)

lemma
/'lemə/, /'lem@/, [N: lemma], [plural: lemmata]. Domain: lexicon. Cohyponym: abstract lemma, lexeme. Def.: 1. A lemma is a lexical access key, for which the canonical orthogrphy of a lexical entry is often used. 2. A headword under which variants of a lexical entry are listed. 3. An abstract lexical unit (abstract lemma).

lexical accent
/'leksɪkəl 'æksənt/, /'leksIk@l '{ks@nt/, [N: [AJ: lexical][N: accent]], [plural: -s]. Hyperonyms: accent. Synonyms: word accent. Cohyponym: phrase accent, sentence accent, syntactical accent, tonal accent. Def.: The emphasis which makes a particular (...) syllable [in a word] stand out (...). (Crystal 1988, p. 2)

lexical category
/'leksɪkəl 'kætəgəri/, /'leksIk@l 'k{t@g@ri/, [N: [AJ: lexical][N: category]], [plural: y/-ies]. Domain: lexicon. Hyperonyms: part of speech. Cohyponym: grammatical category. Def.: Lexical categories are the parts of speech Noun, Adjective, Verb, Adverb, i.e. open classes which may be extended by morphological rules of word formation. (Gibbon et al. 1997, p. 218)

lexical database
/'leksɪkəl 'deɪtəbeɪs/ , /'leksIk@l 'deIt@beIs/ , [N: [AJ: lexical][N: database]], [plural: -s]. Domain: lexicon. Hyperonyms: spoken language lexicon. Cohyponym: system lexicon. Def.: A lexical database is a set of more or less loosely related simpler databases (e.g. pronunciation table, index into a signal annotation file database, stochastic word model, linguistic lexical database with syntactic and semantic information). (Gibbon et al. 1997, p. 185)

lexical information model
/'leksɪkəl ɪnfə'meɪʃən 'mɒdəl/, /'leksIk@l Inf@'meIS@n 'mQd@l/, [N: [AJ: lexical][N: information][N: model]], [plural: -s]. Domain: lexicon. Hyperonyms: model. Def.: A model based on specific types of lexical information and lexical objects.

Google Exhibit 1008
Page 0032

lexical item
/'leksɪkəl 'aɪtəm/, /'leksIk@l 'aIt@m/, [N:[AJ: lexical][N: item]], [plural: -s]. Domain: . Hyponyms: idiom, actual word, heterograph, homograph, homohym, homophone, potential word, synonym . Synonyms: lexical entry. Def.: A unit of language such as a word or an idiom which is inventarised in a lexicon.

lexical morph
/'leksɪkəl 'mɔːf/, /'leksIk@l 'mO:f/, [N: [AJ: lexical][N: morph]], [plural: -s]. Domain: lexicon. Hyperonyms: morph. Hyponyms: phonological lexical morph, grammatical lexical morph. Cohyponym: grammatical morph. Def.: Orthographic or phonological realisation of a lexical morpheme.

lexical morpheme
/'leksɪkəl 'mɔːfiːm/, /'leksIk@l 'mO:fi:m/, [N: [AJ: lexical][N: morpheme]], [plural: -s]. Domain: lexicon. Hyperonyms: morpheme. Cohyponym: grammatical morpheme. Def.: 1. A lexical morpheme is characterised by membership of a large, potentially open class, with meanings such as properties and roles of objects, states and events. (Gibbon et al. 1997, p. 214) 2. Indivisable word stem. (Gibbon et al. 1997, p. 196)

lexical object
/'leksɪkəl 'ɒbdʒekt/, /'leksIk@l 'QbdZekt/, [N: [AJ: lexical][N: object]], [plural: -s]. Domain: lexicon. Hyponyms: lexical sign class, archi-sign; lexicon, lexical database. Def.: 1. The basic object (such as a morpheme, a stem, a word or an idiom) described in a lexicon. 2. The lexical sign class or archi-sign in which similar lexical objects are grouped together, each characterised by subsets of the lexical information required to characterise specific lexical signs. (Gibbon et al. 1997, p. 192)

lexical sign class
/'leksɪkəl 'saɪn 'klɑːs/, /'leksIk@l 'saIn 'klA:s/, [N: [AJ: lexical][N: sign][N: class]], [plural: -es]. Domain: lexicon. Hyperonyms: lexical object. Synonyms: archi-sign. Def.: Group of similar lexical objects, each characterised by subsets of the lexical information required to characterise specific lexical signs. (Gibbon et al. 1997, p. 192)

lexical sign
/'leksɪkəl 'saɪn/, /'leksIk@l 'saIn/, [N: [AJ: lexical][N: sign]], [plural: -s]. Domain: lexicon. Hyperonyms: lexical object. Def.: A sign which is inventarised in a lexicon and whose properties are not (or at least not wholly) composed from those of constituent signs. A lexical sign has properties, often represented as attribute-value pairs, and known as lexical information. The microstructure of a lexicon is the structure associated with the types of lexical information assigned to a lexical sign. (Gibbon et al. 1997, p. 192)

lexical unit
/'leksɪkəl 'juːnɪt/, /'leksIk@l 'ju:nIt/, [N: [AJ: lexical][N: unit]], [plural: -s]. Domain: lexicon. Hyperonyms: speech unit. Hyponyms: word, idiom, discourse particle, hesitation, pragmatic idiom, functional unit. Synonyms: lexical item. Def.: A unit of language such as a word or an idiom which is inventarised in a lexicon.

lexicon architecture
/'leksɪkən 'ɑːkɪtektʃə/, /'leksIk@n 'A:kItektS@/, [N: [N: lexicon][N: architecture]], [plural: -s]. Domain: lexicon. Hyperonyms: structure; linguistic characterisation. Synonyms: lexicon macrostructure. Def.: The choice of basic objects and properties in the lexicon, and the macrostructure of the lexicon as a whole, such as a table of items, a trie (decision tree), an inheritance hierarchy, a semantic network, a database.

Google Exhibit 1008
Page 0033

lexicon formalism
/ˈleksɪkən ˈfɔːməlɪzəm/, /'leksIk@n 'fO:m@lIz@m/, [N: [N: lexicon][N: formalism]], [plural: -s]. Domain: lexicon. Hyperonyms: formal language. Cohyponym: lexicon theory, lexicon model, linguistic framework. Def.: A specially designed logic programming language such as DATR, or an algebraic formalism such as attribute-value matrices, or appropriate definitions in high level languages such as LISP or Prolog, with compiler concepts for translating these languages into conventional languages for efficient processing. (Gibbon et al. 1997, p. 192)

lexicon model
/ˈleksɪkən ˈmɒdəl/, /'leksIk@n 'mQd@l/, [N: [N: lexicon][N: model]], [plural: -s]. Domain: lexicon. Hyperonyms: model; linguistic characterisation. Cohyponym: lexicon theory, lexicon formalism, linguistic framework. Def.: A lexicon model is the specification of the domain denoted by a lexicon theory, conceptually independent of the theory itself. A different definition is also common: the general structure of the objects and attribute-value structures in a formal lexicon. A lexicon model specifies the following kinds of information: - Types of lexical object and structure of lexical entries. - Types of lexical information associated with lexical objects in lexical entries. - Relations between lexical objects and structure of the lexicon as a whole lexicon architecture. (Gibbon et al. 1997, p. 193)

lexicon structure
/ˈleksɪkən ˈstrʌktʃə/, /'leksIk@n 'strVktS@/, [N: [N: lexicon][N: structure]], [plural: -s]. Domain: lexicon. Hyponyms: lexicon macrostructure, lexicon microstructure. Synonyms: lexicon architecture. Def.: The organisation of information in lexica in terms of macrostructure (the overall structure in which lexical entries are located) and microstructure (the internal structure of a lexical entry). (Gibbon et al. 1997, p. 221)

lexicon theory
/ˈleksɪkən ˈθɪəri/, /'leksIk@n 'TI@ri/, [N: [N: lexicon][N: theory]], [plural: y/-ies]. Domain: lexicon. Hyperonyms: theory; linguistic characterisation. Hyponyms: general lexicon theory, specific lexicon theory. Cohyponym: lexicon formalism, lexicon model, linguistic framework. Def.: A coherent and consistent set of expressions formulated in a well-defined formalism and interpreted with respect to a lexicon model. (Gibbon et al. 1997, p. 192)

lexicon
/ˈleksɪkən/, /'leksIk@n/, [N: lexicon], [plural: -s, lexica]. Domain: . Hyperonyms: lexical object, . Hyponyms: declarative lexicon, fully inflected form lexicon, morph lexicon, morpheme lexicon, procedural lexicon, pronunciation lexicon, prosodic lexicon, stem lexicon, . Cohyponym: grammar. Def.: A collection of lexical items such as words or idioms organised within a lexical macrostructure (a list, a tree structure, etc.), each of which has a regular microstructure to which types of lexical information are assigned.

linear discriminant analysis
/ˈlɪnɪə dɪsˈkrɪmɪnənt əˈnæləsɪs/, /'lInI@ dIs'krImIn@nt @'n{l@sIs/, [N: [AJ: linear][AJ: discriminant][N: analysis]], [plural: linear discriminant analyses]. Hyperonyms: statistical technique. Def.: A statistical technique that under certain assumptions finds a linear transformation that will best separate a set of classes when the classification decision is the identification of the closest class centroid measured using Euclidean distances.

linear interpolation
/ˈlɪnɪər ɪntɜːpəˈleɪʃən/, /'lInI@r Int3:p@'leIS@n/, [N: [AJ: linear][N: interpolation]], [plural: -s]. Domain: language modelling. Hyperonyms: interpolation, smoothing technique. Cohyponym: linear discounting, absolute discounting. Def.: 1. Interpolation of unknown values of a linear function y=c+mx between two known values. 2. A technique in the context of language model smoothing by which the relative frequencies of a specific model are interpolated with those of a more general model.

Google Exhibit 1008
Page 0034

MVIP
/'em 'viː 'aɪ 'piː/, /'em 'vi: 'aI 'pi:/, [N: MVIP], [plural: -s]. Domain: system design. Hyperonyms: bus. Synonyms: Multi-Vendor Integration Protocol. Def.: MVIP is a multiplexed digital telephony highway for use within one computer chassis. It provides standard connection for digital telephone traffic between individual circuit boards. It supports telephone circuit-switching under direct computer control, using digital switch elements distributed amongst circuit boards in a standard computer. MVIP software standards allow system integrators to combine MVIP-compatible products from different vendors. The communication technologies that are supported include call management, voice store and forward, speech recognition, text-to-speech, Fax, data communications, and digital circuit-switching. The objective of an MVIP bus is to carry telephone traffic. It allows the interface to the telephone network to be separated from voice processing resources so that the telephone interface may be obtained from one vendor while the voice processing resources are obtained from others. A single MVIP bus has a capacity of 256 full-duplex telephone channels.

n-gram grammar
/'engræm 'græmə/, /'engr{m 'gr{m@/, [N:[N: n-gram][N: grammar]], [plural: -s]. Domain: . Cohyponym: null grammar, unigram grammar, bigram grammar, trigram grammar. Def.: A probabilistic grammar based on transition probabilities of words, predicting the probability of a word in a given context from the product of the a priori probability of the word and the probability of its (n-1) predecessors. The transition probabilities of words are calculated from their distribution in a corpus. Analogously, the probability of a word in a bigram or trigram grammar is calculated using the probabilities of the preceding word or the preceding two words.

n-gram language model
/'engræm 'læŋgwɪdʒ 'mɒdəl/, /'engr{m 'l{NgwIdZ 'mQd@l/, [N: [N: n-gram][N: language][N: model]], [plural: -s]. Domain: language modelling. Hyperonyms: stochastic language model. Synonyms: n-gram language model. Cohyponym: bigram language model, trigram language model. Def.: An n-gram language model is a stochastic language model that is based on a n-gram grammar, i.e. on conditional probabilities depending only on the (n-1) immediate predecessor words. (Gibbon et al. 1997, p. 243)

narrow phonetic transcription
/'nærəʊ fə'netɪk træn'skrɪpʃən/, /'n{r@U f@'netIk tr{n'skrIpS@n/, [N: [AJ: narrow][AJ: phonetic][N: transcription]], [plural: -s]. Domain: corpora, lexicon. Hyperonyms: phonetic transcription. Cohyponym: broad phonetic transcription. Def.: Phonetic transcription which is relatively detailed and represents more than the minimum number of phonetic features which are needed to represent phonemes. (Crystal 1988, p. 313)

nasal
/'neɪzəl/, /'neIz@l/, [N: nasal], [plural: -s]. Hyperonyms: consonant; manner of articulation. Cohyponym: plosive, trill, tap, flap, fricative, lateral fricative, approximant, lateral approximant. Def.: A nasal consonant or vowel is classified on the basis of manner of articulation: a sound produced while the soft palate is lowered to allow resonance of the nasal cavity and escape of air through the nose. (Crystal 1988, p. 203) E.g. [m, n, N].

Natural Language Processing
/'nætʃərəl 'læŋgwɪdʒ 'prəʊsesɪŋ/, /'n{tS@r@l 'l{NgwIdZ 'pr@UsesIN/, [N: [AJ: Natural][N: Language][N: Processing]], [plural: none]. Synonyms: NLP. Cohyponym: SLP, spoken language processing. Def.: The parsing of written language into semantic concepts.

natural language
/'nætʃərəl 'læŋgwɪdʒ/, /'n{tS@r@l 'l{NgwIdZ/, [N: [AJ: natural][N: language]], [plural: -s]. Hyperonyms: language. Cohyponym: formal language. Def.: The forms of language used by humans for communication, as opposed to artificial or formal languages such as logics, algebras, programming languages. The controlled language used between people and invented systems can be also termed "natural" if it is what users spontaneously produce in response to the situation.

Google Exhibit 1008
Page 0035

near end echo
/nɪər'end 'ekəʊ/, /nI@r'end 'ek@U/, [N: [AJ: near][N: end][N: echo]], [plural: -es]. Hyperonyms: echo. Cohyponym: far end echo. Def.: The electrical echo of the output of a telephone system generated at the point where the system is connected to the network. Near end echo is always an electrical phenomenon.

neural network based approach
/'njuːrəl 'netwɜːk 'beɪst ə'prəʊtʃ/, /'nju:r@l 'netw3:k 'beIst @'pr@UtS/, [N: [AJ: neural][N: network][AJ:based][N: approach]], [plural: -es]. Domain: multimodal systems. Hyperonyms: template matching. Cohyponym: Principal Component Analysis, PCA, geometric template matching, deformable template matching, optical flow technique. Def.: Another variant of template matching are neural-network based approaches to, for instance, speech or face recognition, for example applying multi-layer perceptrons or Kohonen self-organising maps.

neural network
/'njuːrəl 'netwɜːk/, /'nju:r@l 'netw3:k/, [N: [AJ: neural][N: network]], [plural: -s]. Hyponyms: multi-layer perceptrion, back-propagation network, Kohonen map. Def.: An artificial neural network is a pattern recognition system modelled on the human nervous system, in which input values are mapped via a network of weighted functions (nodes) into output classes.

newsgroup
/'njuːzgruːp/, /'nju:zgru:p/, [N: newsgroup], [plural: -s]. Def.: Moderated or unmoderated Internet information service; a newsgroup is focused on a specific subject, e.g. Spoken Language Processing (SLP). Users subscribe to a newsgroup to read and post mailings.

NIST-SPHERE header
/'nɪst 'sfiːə 'hedə/, /'nIst 'sfi:@ 'hed@/, [N: [N: NIST-SPHERE][N: header]], [plural: -s]. Hyperonyms: audio signal header. Def.: Standard header for audio signals proposed by the National Institute of Standards and Technology (NIST), consisting of a simple ASCII formatted text information describing the signal following the header; the header options can be user-specified.

NLP
/'en 'el 'piː/, /'en 'el 'pi:/, [N: NLP], [plural: none]. Synonyms: Natural Language Processing. Cohyponym: SLP, Spoken Language Processing. Def.: A theoretical and applied discipline, involving computational linguistics and artificial intelligence, which is concerned with natural language understanding (parsing and interpretation) and natural language generation (production of sentences and texts).

noise-cancelling microphone
/nɔɪz 'kænsəlɪŋ 'maɪkrəfəʊn/, /nOIz 'k{ns@lIN 'maIkr@f@Un/, [N: [N: noise][V: cancelling][N: microphone]], [plural: -s]. Hyperonyms: microphone. Synonyms: close-talking microphone. Def.: A noise-cancelling microphone is more sensitive to nearby sound sources (i.e. the mouth) than to distant sound sources (e.g. noise sources).

non-adaptive language model
/nɒnə'dæptɪv 'læŋgwɪdʒ 'mɒdəl/, /nQn@'d{ptIv 'l{NgwIdZ 'mQd@l/, [N: [AJ: non-adaptive][N: language][N: model]], [plural: -s]. Domain: language modelling. Hyperonyms: language model. Cohyponym: adaptive language model. Def.: A non-adaptive language model does not depend on the test data, but remains unchanged as trained on the original training data. (Gibbon et al. 1997, p. 257)

Google Exhibit 1008
Page 0036

ontological hierarchy
/ɒntəˈlɒdʒɪkəl ˈhaɪərɑːki/, /Qnt@'lQdZIk@l 'haI@rA:ki/, [N: [AJ: ontological][N: hierarchy]], [plural: y/-ies]. Domain: terminology. Hyperonyms: hierarchy. Synonyms: partitive hierarchy, meronomy, mereonomy, PARTOF hierarchy. Cohyponym: taxonomy, ISA hierarchy, logical concept hierarchy, generic concept hierarchy. Def.: Hierarchy of concepts in a PARTOF relation, i.e. a mereonomy.

OOV word
/ˈəʊ ˈəʊ ˈviː ˈwɜːd/, /'@U '@U 'vi: 'w3:d/, [N: [N: OOV][N: word]], [plural: -s]. Domain: speech recognition, language modelling. Synonyms: out-of-vocabulary word. Def.: Word not listed in the lexicon of a spoken language system.

OOV-rejection
/ˈəʊ ˈəʊ ˈviː rɪˈdʒekʃən/, /'@U '@U 'vi: rI'dZekS@n/, [N: [N: OOV][N: rejection]], [plural: -s]. Domain: speech recognition, consumer off-the-shelf products. Hyperonyms: performance measure. Cohyponym: recognition accuracy, error recovery, response time, situational awareness. Def.: Rejection rate for out-of-vocabulary words.

opinion testing
/əˈpɪnɪən ˈtestɪŋ/ , /@'pInI@n 'testIN/ , [N: [N: opinion][N: testing]] , [plural: none] . Domain: speech synthesis. Hyperonyms: testing procedure. Synonyms: judgment testing. Cohyponym: functional testing. Def.: Procedure whereby a group of listeners is asked to judge the performance of a speech output system along a number of rating scales.

optical flow technique
/ˈɒptɪkəl ˈfləʊ tekˈniːk/, /'QptIk@l 'flOU tek'ni:k/, [N: [AJ: optical][N: flow][N: technique]], [plural: -s]. Domain: multimodal systems. Hyperonyms: template matching. Cohyponym: Principal Component Analysis, PCA, geometric template matching, deformable template matching, neural network based approach. Def.: The optical flow technique allows the detection of motion (from a sequence of images) rather than facial feature displacements, by computing the difference in image intensity between two consecutive frames. It works at the pixel level. The computation is done pixel per pixel. This technique may be used to extract muscle contraction. Windows are placed around muscle locations. The velocity of each muscle contraction is computed.

oral dialogue
/ˈɔːrəl ˈdaɪəlɒg/, /'O:r@l 'daI@lQg/, [N: [AJ: oral][N: dialogue]] , [plural: -s] . Hyperonyms: dialogue. Def.: See spoken language dialogue. This term is quite widely used, though it is less favoured by native speakers of English than by those who have learned it as a second-language.

organic speech disorder
/ɔːˈgænɪk ˈspiːtʃ dɪsˈɔːdə/, /O:'g{nIk 'spi:tS dIs'O:d@/, [N: [AJ: organic][N: speech][N: disorder]], [plural: -s]. Domain: corpora. Hyperonyms: speech disorder. Cohyponym: functional speech disorder. Def.: Speech disorders where there is a clear organic (anatomical, physiological, neurological) cause. (Gibbon et al. 1997, p. 114)

orthographic break
/ɔːθəˈgræfɪk ˈbreɪk/, /O:T@'gr{fIk 'breIk/, [N: [AJ: orthographic][N: break]], [plural: -s]. Domain: lexicon. Synonyms: orthographic syllable. Def.: The orthographic break is in general only needed for splitting words at line-breaks and does not correspond closely to either syllable or morph boundaries but combines phonological, morphological, orthographic or other criteria. Conventions differ from language to language. (Gibbon et al. 1997, p. 213)

orthographic lexical morph
/ɔːθəˈgræfɪk ˈleksɪkəl ˈmɔːf/, /O:T@'gr{fIk 'leksIk@l 'mO:f/, [N: [AJ: orthographic][AJ: lexical][N: morph]], [plural: -s]. Domain: lexicon. Hyperonyms: lexical morph, orthographic morph. Cohyponym: phonological lexical morph. Def.: The orthographic realisation of a lexical morpheme. (Gibbon et al. 1997, p. 199)

orthographic morph
/ɔːθəˈgræfɪk ˈmɔːf/, /O:T@'gr{fIk 'mO:f/, [N: [AJ: orthographic][N: morph]], [plural: -s]. Domain: lexicon. Hyperonyms: morph. Hyponyms: orthographic lexical morph. Cohyponym: phonological morph. Def.: Orthographic morphs are morphs consisting of graphemes (either single letters or fixed combinations of letters). (Gibbon et al. 1997, p. 215) E.g. 'catches' = catch + es: 'es' is an orthographic morph of the English 3rd P.Sg. Pres. Tense morpheme..

orthographic noise
/ɔːθəˈgræfɪk ˈnɔɪz/, /O:T@'gr{fIk 'nOIz/, [N: [AJ: orthographic][N: noise]], [plural: none]. Domain: lexicon. Def.: Intrusive artefacts in text-to-speech or speech recognition systems due to the use of orthography in interface specifications; orthographic noise is due to homography, homophony, and other irregularities in grapheme-phoneme relationships at character and word level. (Gibbon et al. 1997, p. 203)

orthographic transcription
/ɔːθəˈgræfɪk trænˈskrɪpʃən/, /O:T@'gr{fIk tr{n'skrIpS@n/, [N: [AJ: orthographic][N: transcription]], [plural: -s]. Hyperonyms: transcription . Cohyponym: phonemic transcription, phonetic transcription. Def.: Transcription of a speech signal in standard orthography.

orthographic word
/ɔːθəˈgræfɪk ˈwɜːd/, /O:T@'gr{fIk 'w3:d/, [N: [AJ: orthographic][N: word]], [plural: -s]. Domain: lexicon. Hyperonyms: word. Cohyponym: graphemic word, phonetic word, phonological word, morphological word, syntactic word, prosodic word. Def.: A word defined in terms of its conformity to the orthographic structure and punctuation of a language, for instance as delimited by spaces or punctuation marks. (Gibbon et al. 1997, p. 196)

orthography
/ɔːˈθɒgrəfi/, /O:'TQgr@fi/, [N: orthography], [plural: y/-ies]. Domain: lexicon, graphemics. Hyperonyms: surface representation. Hyponyms: alphabetic orthography, syllabic orthography, logographic orthography; Roman orthography, Cyrillic orthography, Greek orthography; standard orthography, reformed orthography. Synonyms: spelling. Cohyponym: pronunciation, phonology, phonetics. Meronym. sup.: writing. Meronym. sub.: letters, characters. Def.: The two-dimensional visual representation of word forms by standardised graphical characters (Gibbon et al. 1997, p. 188)

OSQL
/ˈəʊ ˈes ˈkjuː ˈel/, /'@U 'es 'kju: 'el/, [N: OSQL], [plural: none]. Hyperonyms: query language. Cohyponym: SQL, Standard Query Language. Def.: An object-oriented high-level database query language for which a draft standard has been published by the ODMG.

out-of-vocabulary word
/ˈaʊt əv vəˈkæbjʊləri ˈwɜːd/, /'aUt @v v@'k{bjUl@ri 'w3:d/, [N: [PREP: out][PREP: of][N: vocabulary][N: word]], [plural: -s]. Domain: speech recognition, language modelling . Synonyms: OOV word. Def.: Word not listed in the lexicon of a spoken language system.

output device
/ˈaʊtpʊt dɪˈvaɪs/, /'aUtpUt dI'vaIs/, [N: [N: output][N: device]], [plural: -s]. Domain: multimodal systems. Hyperonyms: technical device. Hyponyms: visual device, acoustic device, haptic device. Def.: Modality used by a multimodal system to communicate with the user.

output medium
/ˈaʊtpʊt ˈmiːdɪəm/, /'aUtpUt 'mi:dI@m/, [N: [N: output][N: medium]], [plural: um/-a]. Domain: multimodal systems. Hyperonyms: modality. Hyponyms: non-speech output modality. Cohyponym: input modality. Def.: An output channel from machine to human in human-machine communication. Output such as synthesised speech, synthesised faces, talking heads (combination of speech and face synthesis), synthetic agents, force feedback, traditional multimedia output (text, graphics, video, sound).

Google Exhibit 1008
Page 0038

output modality
/ˈaʊtpʊt məˈdælɪti/, /'aUtpUt m@'d{lIti/, [N: [N: output][N: modality]], [plural: -s]. Domain: multimodal systems. Hyperonyms: modality. Hyponyms: non-speech output modality. Cohyponym: input modality. Def.: A human speech or gestural output channel used in communication, generally acoustic or visual.

paired comparison
/ˈpeəd kəmˈpærɪsən/, /'pe@d k@m'p{rIs@n/, [N: [AJ: paired][N: comparison]], [plural: -s]. Domain: speech synthesis, assessment methodologies. Hyperonyms: testing procedure. Def.: A psychophysical testing procedure that is used when subjects are required to judge between two stimuli. In Language Engineering this might be judging which of two recogniser outputs has more or less intelligibility.

palatal consonant
/ˈpælətəl ˈkɒnsənənt/, /'p{l@t@l 'kQns@n@nt/, [N: [AJ: palatal][N: consonant]], [plural: -s]. Hyperonyms: consonant. Synonyms: palatal. Cohyponym: bilabial consonant, labiodental consonant, dental consonant, alveolar consonant, postalveolar consonant, retroflex consonant, velar consonant, uvular consonant, pharyngeal consonant, glottal consonant. Def.: Palatal consonant is a term used in the phonetic classification of speech sounds on the basis of their place of articulation: it refers to a sound made when the front of the tongue is in contact with or approaches the hard palate. (Crystal 1988, p. 219)

paradigmatic relation
/pærədɪgˈmætɪk rɪˈleɪʃən/, /p{r@dIg'm{tIk rI'leIS@n/, [N: [AJ: paradigmatic][N: relation]] , [plural: -s]. Domain: lexicon. Hyperonyms: linguistic characterisation. Cohyponym: syntagmatic relation. Def.: Paradigmatic relations are classificatory, disjunctive, element-class, subclass-superclass relations.

PARADISE
/ˈpærədaɪs/, /'p{r@daIs/, [N: PARADISE], [plural: none]. Domain: multimodal systems. Hyperonyms: evaluation framework. Def.: Framework for evaluating dialogue systems from a user point of view. It assumes that the ultimate measure of success for a dialogue system is user satisfaction. Since many different factors influence user satisfaction, dependent on the application, PARADISE proposes to use statistical methods of determining the most significant predictors of cumulative user satisfaction for a specific application, out of a large set of potentially useful variables.

paralinguistic feature
pærəlɪŋˈgwɪstɪk ˈfiːtʃə/, p{r@lIN'gwIstIk 'fi:tS@/, [N: [AJ: paralinguistic][N: feature]], [plural: -s]. Domain: dialogue representation. Def.: Concomitant properties of voice such as laughter, tempo, loudness, and so on that occur during speech. We exclude features that do not accompany speech but rather occur in isolation.

parametric model
/pærəˈmetrɪk ˈmɒdəl/, /p{r@'metrIk 'mQd@l/, [N: [AJ: parametric][N: model]], [plural: -s]. Domain: multimodal systems. Hyperonyms: synthetic model. Def.: A facial model is created and animated through a set of parameters. Generally, parameters can be divided into two groups: conformation and expression parameters. The former refer to parameters acting on the facial topology (including position and size of the nose and eyes, global size of the head). The latter specify facial expressions such as brow action, mouth movement, and blink.

parametric synthesis
/pærəˈmetrɪk ˈsɪnθəsɪs/, /p{r@'metrIk 'sInT@sIs/, [N: [AJ: parametric][N: synthesis]], [plural: parametric syntheses]. Domain: speech synthesis. Hyperonyms: speech synthesis. Cohyponym: concatenative synthesis. Def.: Speech synthesis by modelling the (human) articulatory or vocal tract. Parameters control the shape of the vocal tract and hence determine the speech output.

Google Exhibit 1008
Page 0039

parametric test
/pærə'metrɪk test/, /p{r@'metrIk test/, [N: [AJ: parametric][N: test]], [plural: -s]. Domain: assessment methodologies. Hyperonyms: testing procedure. Cohyponym: non-parametric test, distribution-free test. Def.: Statistical test employed in simple hypothesis testing: parametric tests are used when continuous measures are available.

parsing
/'pɑːzɪŋ/, /'pA:zIN/, [N: parsing], [plural: none]. Domain: speech recognition, speech synthesis. Hyponyms: morphological parsing, syntactic parsing, prosodic parsing; sentence parsing. Def.: The labelling of the parts of speech or grammatical elements of sentences, e.g. subject, predicate, past tense, noun, verb, either by a human analyst or by means of a parser (a programme based on a parsing algorithm). (Crystal 1988, p. 221)

part of speech
/'pɑːt əv 'spiːtʃ/, /'pA:t @v 'spi:tS/, [N: [N: part][PREP: of][N: speech]], [plural: parts of speech]. Domain: lexicon. Hyponyms: grammatical category, lexical category. Synonyms: syntactic category, POS. Def.: The traditional term for a grammatical class of words. (Crystal 1988, p. 222) E.g. noun, adjective, article, pronoun, verb, adverb, preposition, conjunction, interjection, proper noun, common noun, intransitive verb, transitive verb, ditransitive verb, prepositional verb .

partial action frame
/'pɑːʃəl 'ækʃən 'freɪm/, /'pA:S@l '{kS@n 'freIm/, [N: [AJ: partial][N: action][N: frame]], [plural: -s]. Domain: multimodal systems. Def.: Input from each modality is interpreted separately and then parsed and transformed into semantic frames containing slots that specify command parameters (parameter slots). The information in these (partial) action frames may be incomplete or ambiguous if not all elements of the command were expressed in a single modality. A domain-independent frame-merging algorithm combines partial frames into complete frames.

partial synonym
/'pɑːʃəl 'sɪnənɪm/, /'pA:S@l 'sIn@nIm/, [N: [AJ: partial][N: synonym]], [plural: -s]. Domain: lexicon. Hyperonyms: synonym. Cohyponym: full synonym. Def.: Two words are partial synonyms if they have at least one meaning in common and if either has additional readings not shared by the other. (Gibbon et al. 1997, p. 850) E.g. 'Manual' and 'handbook' are partial synonyms. 'Manual' is also, among other things, a term for a traditional organ keyboard..

partitive hierarchy
/'pɑːtɪtɪv 'haɪərɑːki/, /'pA:tItIv 'haI@rA:ki/, [N: [AJ: partitive][N: hierarchy]], [plural: y/-ies]. Domain: terminology. Hyperonyms: hierarchy. Synonyms: ontological hierarchy, meronomy, mereonomy, PARTOF hierarchy. Cohyponym: taxonomy, ISA hierarchy. Def.: Hierarchy of concepts in a PARTOF relation, i.e. a mereonomy.

PARTOF hierarchy
/'pɑːtəv 'haɪərɑːki/, /'pA:t@v 'haI@rA:ki/, [N: [AJ: PARTOF][N: hierarchy]], [plural: y/-ies]. Domain: terminology. Hyperonyms: hierarchy. Synonyms: meronomy, mereonomy, partitive hierarchy, ontological hierarchy. Cohyponym: taxonomy, ISA hierarchy, logical concept hierarchy, generic concept hierarchy. Def.: A hierarchy defined by the relation of parts to wholes, and parts to parts.

Google Exhibit 1008
Page 0040

perseverative coarticulation
/pə'sevərətɪv kəʊɑːtɪkjuːleɪʃən/, **/p@'sev@r@tIv k@UA:tIkju:leIS@n/**, [N: [AJ: perseverative][N: coarticulation]], [plural: none]. Domain: multimodal systems. Hyperonyms: coarticulation. Synonyms: forward coarticulation, left-to-right coarticulation. Cohyponym: backward coarticulation, anticipatory coarticulation, right-to-left coarticulation. Def.: In the string ...AB..., sound A influences sound B (or beyond). L > R coarticulation is thought to be largely due to lag in articulatory movement, induced by inertia. (Clark & Yallop, p. 87)

personal-password speaker recognition system
/'pɜːsənəl 'pɑːswɜːd 'spiːkə rekəg'nɪʃən 'sɪstəm/, **/'p3:s@n@l 'pA:sw3:d 'spi:k@ rek@g'nIS@n 'sIst@m/**, [N: [AJ: personal][N: password][N: speaker][N: recognition][N: system]], [plural: -s]. Domain: speaker recognition. Hyperonyms: speaker recognition system. Cohyponym: common-password speaker recognition system. Def.: A text-dependent speaker recognition system for which each registered speaker has his own voice password.

pharyngeal consonant
/fə'rɪn'dʒɪəl 'kɒnsənənt/, **/f@'rIn'dZI@l 'kQns@n@nt/**, [N: [AJ: pharyngeal][N: consonant]], [plural: -s]. Hyperonyms: consonant. Cohyponym: bilabial consonant, labiodental consonant, dental consonant, alveolar consonant, postalveolar consonant, retroflex consonant, palatal consonant, velar consonant, uvular consonant, glottal consonant. Def.: Pharyngeal consonant is a term used in the phonetic classification of consonant sounds on the basis of their place of articulation: it refers to a sound made in the pharynx. (Crystal 1988, p. 226)

pharynx
/'færɪŋks/, **/'f{rINks/**, [N: pharynx], [pural: -es]. Hyperonyms: articulator. Meronym. sup.: vocal tract. Def.: Pharynx is the tubular cavity which constitutes the throat above the larynx. (Crystal 1988, p. 226)

phase
/'feɪz/, **/'feIz/**, [N: phase], [plural: -s]. Domain: physical characterisation. Hyperonyms: acoustic measure. Def.: The time displacement between two sinusoidal waveforms of the same frequency. (Clark & Yallop, p. 213)

phone
/'fəʊn/, **/'f@Un/**, [N: phone], [plural: -s]. Meronym. sup.: word. Def.: 1. A segment of a spoken utterance which is assigned to a single phoneme; in this role it is called an allophone of that phoneme. 2. Informally, in speech technology, a subword unit of speech that represents a particular sound.

phoneme monitoring
/'fəʊniːm 'mɒnɪtərɪŋ/ , **/'f@Uni:m 'mQnIt@rIN/** , [N: [N: phoneme][N: monitoring]], [plural: none]. Domain: speech synthesis. Hyperonyms: monitoring. Cohyponym: word monitoring, syllable monitoring. Def.: Testing the intelligibility of individual sounds. (Gibbon et al. 1997, p. 490)

phoneme
/'fəʊniːm/, **/'f@Uni:m/**, [N: phoneme], [plural: -s]. Hyperonyms: phonological unit. Def.: 1. In traditional phonology: the smallest distinctive unit of sound in a language, where 'distinctive' means 'which distinguishes one word from another'; phoneme contrasts are often tested with minimal pairs, i.e. words differing only in one phoneme at the same position, e.g. /miːl/ - /piːl/ 'meal' - 'peal'. 2. In structuralist phonology: A set of phonetically similar sounds in complementary distribution (allophones). 3. In Prague and generative phonology: a bundle of distinctive features. The phoneme is the reference unit for alphabetic orthographies. Because of theoretical and practical problems with the notion of phoneme (e.g. coarticulation between adjacent phonemes, highly restricted distributional patterns within syllables, the importance of hierarchically larger units) phonological developments in the last quarter of the 20th century have not taken the phoneme to be a basic unit of linguistic description but rather an epiphenomenon. (Gibbon et al., p. 212)

phonemic transcription
/fə'ni:mɪk træn'skrɪpʃən/, /f@'ni:mIk tr{n'skrIpS@n/, [N: [AJ: phonemic][N: transcription]], [plural: -s]. Hyperonyms: transcription. Synonyms: broad phonetic transcription. Cohyponym: phonetic transcription, orthographic transcription. Def.: A transcription in which only phonemes, i.e. minimal contrastive units in the sound system of a language, are used, and non-minimal phonetic information is excluded. Citations of phonemic transcriptions are conventionally enclosed in oblique bars (slashes), i.e. /.../. E.g. /pIn/, /pen/, /p{n/.

phonetic transcription
/fəʊ'netɪk træn'skrɪpʃən/, /f@U'netIk tr{n'skrIpS@n/, [N: [AJ: phonetic][N: transcription]], [plural: -s]. Hyperonyms: transcription. Hyponyms: broad phonetic transcription, narrow phonetic transcription. Cohyponym: phonemic transcription. Def.: A phonetic transcription gives details of pronunciation beyond the phonemically minimal features. The relation between the phonemic and the phonetic level can be described by general rules mapping phonemes to their detailed realisations (allophones) in specific contexts. Citations of phonetic forms are standardly delimited by square brackets [...]. (Gibbon et al. 1997, p. 209) E.g. German [?apf@l] 'apple'.

phonetically balanced sentence
/fə'netɪkəli 'bælənst 'sentəns/, /f@'netIk@li 'b{l@nst 'sent@ns/, [N: [AV: phonetically][AJ: balanced][N: sentence]], [plural: -s]. Domain: corpora, assessment methodologies, speaker recognition, system design. Hyperonyms: sentence. Cohyponym: phonetically rich sentence. Def.: Sentence containing phonemes according to their frequency of occurrence in a given language.

phonetically rich sentence
/fə'netɪkəli rɪtʃ 'sentəns/, /f@'netIk@li rItS 'sent@ns/, [N: [AV: phonetically][AJ: rich][N: sentence]] , [plural: -s]. Domain: corpora. Hyperonyms: sentence. Cohyponym: phonetically balanced sentence. Def.: Sentence containing approximately uniform phoneme frequency distributions.

phonetics
/fə'netɪks/, /f@'netIks/, [N: phonetics], [plural: always plural]. Meronym. sup.: linguistics. Meronym. sub.: articulatory phonetics, acoustic phonetics, auditory phonetics. Def.: The science of the sounds of human languages, their production, transmission and perception, whose methodology encompasses both the expert judgment and transcription of sound properties by a phonetician and the quantitative physical measurement and statistical evaluation of the events involved in articulation, transmission and perception. (Crystal 1988, p. 229)

phonological lexical morph
/fɒnə'lɒdʒɪkəl 'leksɪkəl 'mɔ:f/, /fQn@'lQdZIk@l 'leksIk@l 'mO:f/, [N: [AJ: phonological][AJ: lexical][N: morph]], [plural: -s]. Domain: lexicon. Hyperonyms: lexical morph, phonological morph. Cohyponym: orthographic lexical morph. Def.: The phonological realisation of a lexical morpheme. (Gibbon et al. 1997, p. 199)

phonological morph
/fɒnə'lɒdʒɪkəl 'mɔ:f/, /fQn@'lQdZIk@l 'mO:f/, [N: [AJ: phonological][N: morph]], [plural: -s]. Domain: lexicon. Hyperonyms: morph. Cohyponym: orthographic morph. Def.: In traditional phonology: phonological morphs are morphs consisting of phoneme sequences with a prosodic pattern (e.g. word stress). (Gibbon et al. 1997, p.215) E.g. recognition /,rek@g'nIS@n/.

phonological unit
/fɒnə'lɒdʒɪkəl 'ju:nɪt/, /fQn@'lQdZIk@l 'ju:nIt/, [N:[AJ: phonological][N: unit]], [plural: -s]. Hyponyms: phoneme syllable. Def.: A unit of phonological description such as the distinctive feature, the phoneme, the syllable.

Google Exhibit 1008
Page 0042

Principal Component Analysis
/ˈprɪnsɪpəl kəmˈpəʊnənt əˈnæləsɪs/, /'prInsIp@l k@m'p@Un@nt @'n{l@sIs/, [N: [AJ: Principal][N: Component][N: Analysis]], [plural: none]. Domain: multimodal systems. Hyperonyms: template matching. Synonyms: PCA. Cohyponym: geometric template matching, deformable template matching, optical flow technique, neural network based approach. Def.: The simplest version of template matching. A test image is classified based on its (Euclidean) distance to templates generated from the faces in the training set (database). The Kurhunen-Loeve procedure and eigenfaces are based on this simple template matching method. Eigenfaces correspond to characteristic feature images and can be viewed as the principal components of a test image with respect to characteristic features obtained from the database of faces. This technique has been applied to recognise lip shapes.

prior probability
/ˈpraɪə prɒbəˈbɪlɪti/, /'praI@ prQb@'bIlIti/, [N: [AJ: prior][N: probability]], [plural: y/-ies]. Domain: language modelling. Hyperonyms: probability. Synonyms: a priori probability, priors. Cohyponym: posterior probability. Def.: The probability of observing some phenomenon estimated from previously collected training data, but independent of future observations.

procedural model
/prəˈsiːdʒərəl ˈmɒdəl/, /pr@'si:dZ@r@l 'mQd@l/, [N:[AJ: procedural][N: model]], [plural: -s]. Domain: multimodal systems. Hyperonyms: animation control technique, synthetic model. Cohyponym: parametric model, free form deformation. Def.: This method is not based on biological studies. Rather, the idea is to simulate the action of a muscle by a few parameters. Muscles are simulated by specialised procedures. These procedures are considered as Abstract Muscle Actions (AMAs) and can have up to 24 parameters. They work on specific facial regions that correspond to one muscle. They compute the displacement occurring under muscle contraction.

production model
/prəˈdʌkʃən ˈmɒdəl/, /pr@'dVkS@n 'mQd@l/, [N: [N: production][N: model]], [plural: -s]. Domain: speech synthesis, consumer off-the-shelf products. Hyperonyms: speech synthesis model. Cohyponym: concatenation technique, playback technique. Def.: A physical model of the vocal folds and the vocal tract used to produce sounds that resemble speech. These models are often LPC based (specifying sounds by formants in position and width). The voice quality is not as good as in the playback or concatenation technique, but since this technique comprises pure synthesis, every parameter is controlled. Thus the change of voice characteristics, pitch, intonation, and stress can be exploited by the system. The vocabulary is limited by the pronunciation rules.

pronunciation lexicon
/prənʌnsiˈeɪʃən ˈleksɪkən/ , /pr@nVnsi'eIS@n 'leksIk@n/ , [N: [N: pronunciation][N: lexicon]], [plural: pronunciation lexica ,-s]. Domain: lexicon. Hyperonyms: lexicon. Synonyms: pronunciation dictionary, pronunciation table. Meronym. sup.: acoustic-phonetic model. Def.: A table containing pairs of orthographic and phonemic representations of words, sometimes with variant orthographies and pronunciations, and either in book form or as a database.

pronunciation table
/prənʌnsiˈeɪʃən ˈteɪbəl/, /pr@nVnsi'eIS@n 'teIb@l/, [N: [N: pronunciation][N: table]], [plural: -s]. Domain: lexicon. Hyperonyms: database type. Synonyms: pronunciation dictionary, pronunciation lexicon. Def.: A pronunciation table defines the relation between orthographic and phonemic representations of words. Often they are defined as functions which assign pronunciations (frequently a set of variant pronunciations) to orthographic representations (Gibbon et al. 1997, p. 226)

prosodic feature
/prəˈzɒdɪk ˈfiːtʃə/, /pr@'zQdIk 'fi:tS@/, [N: [AJ: prosodic][N: feature]], [plural: -s]. Hyponyms: pitch, loudness, tempo, rhythm. Def.: A term used in suprasegmental phonetics and phonology to refer collectively to variations in pitch, loudness, tempo and rhythm. (Crystal 1988, p. 249)

Google Exhibit 1008
Page 0043

prosodic lexicon
/prə'zɒdɪk 'leksɪkən/, /pr@'zQdIk 'leksIk@n/, [N: [AJ: prosodic][N: lexicon]], [plural: prosodic lexica, -s]. Domain: lexicon. Hyperonyms: lexicon. Def.: A lexicon of prosodic units such as terminal pitch contours (rises, falls, fall-rises), with their meanings.

prosodic parsing
/prə'zɒdɪk 'pɑːzɪŋ/, /pr@'zQdIk 'pA:zIN/, [N: [AJ: prosodic][N: parsing]], [plural: none]. Domain: speech recognition, speech synthesis. Hyperonyms: parsing. Cohyponym: morphological parsing, syntactic parsing. Def.: 1. In speech recognition: Prosodic parsing is the analysis of the speech signal in respect of the fundamental frequency (F0) trajectory in relation to words, sentences and dialogue units. 2. In speech synthesis: Prosodic parsing is the analysis of sentence structure for the generation of intonation patterns in speech synthesis. (Gibbon et al. 1997, p. 210)

prosodic transcription
/prə'zɒdɪk træn'skrɪpʃən/, /pr@'zQdIk tr{n'skrIpS@n/, [N: [AJ: prosodic][N: transcription]], [plural: -s]. Hyperonyms: transcription. Def.: A transcription of the prosodic features of an utterance such as stress/accent, intonation, boundaries, using a prosodic transcription system or alphabet such as ToBI or SAMPROSA.

prosodic word
/prə'sɒdɪk 'wɜːd/, /pr@'sQdIk 'w3:d/, [N: [AJ: prosodic][N: word]], [plural: -s]. Domain: lexicon. Hyperonyms: word. Cohyponym: orthographic word, phonological word, morphological word, syntactic word. Def.: Word based on its conformity to the accentuation and the rhythm patterning of the language. (Gibbon et al. 1997, p. 196)

prosody
/'prɒsədi/, /'prQs@di/, [N: prosody], [plural: –]. Synonyms: non-segmental phonology. Meronym. sup.: phonology. Def.: Prosody covers all properties of pronunciation which are not directly concerned with defining consonants and vowels. Prosody in this sense covers, for example, syllable structure and phonological word phonotactics, as well as the more traditional categories of intonation, accent, and phrasing. In the lexicon, prosodic information is in general restricted to the prosodic properties of words, such as stress position (e.g. in English, Dutch, and German words), or tonal accent words (e.g. in Swedish), or to rhythmically relevant units such as the syllable and the foot. (Gibbon et al. 1997, p. 210-211)

pseudo-impostor bundle
/'sjuːdəʊ ɪm'pɒstə 'bʌndəl/, /'sju:d@U Im'pQst@ 'bVnd@l/, [N: [AJ: pseudo][N: impostor][N: bundle]], [plural: -s]. Domain: speaker recognition. Hyperonyms: speaker group. Def.: The group of speakers who have been used to build the impostor model of a given registered speaker. (Gibbon et al. 1997, p. 421)

pseudo-impostor
/'sjuːdəʊ ɪm'pɒstə/, /'sju:d@U Im'pQst@/, [N: [AJ: pseudo][N: impostor]], [plural: -s]. Domain: speaker recognition. Hyperonyms: impostor. Synonyms: background speaker. Def.: Speaker used to model the impostor during the registration phase. (Gibbon et al. 1997, p. 421)

psycholinguistics
/saɪkəʊlɪŋ'gwɪstɪks/, /saIk@UlIN'gwIstIks/, [N: psycholinguistics], [plural: always plural]. Meronym. sup.: linguistics. Def.: The study of the mental processes underlying the planning, production, perception and comprehension of speech. (Crystal 1988, p. 251)

Pulse coded modulation Expansion Bus
/p'ʌls 'kəʊdɪd mɒdjʊ'leɪʃən ɪks'pænʃən 'bʌs/, /p'Vls 'k@UdId mQdjU'leIS@n Iks'p{nS@n 'bVs/, [N: [N: Pulse][AJ: coded][N: modulation][N: Expansion] [N: Bus]], [plural: -es]. Domain: system design. Hyperonyms: bus. Synonyms: PEB. Def.: PEB is seen as an internal switching matrix capable of routing any time slot to an adequate audio port of the speech recogniser.

Google Exhibit 1008
Page 0044

recognition component
/rekəgˈnɪʃən kəmˈpəʊnənt/ , /rek@g'nIS@n k@m'p@Un@nt/ , [N: [N: recognition][N: component]], [plural: -s]. Domain: speech recognition. Cohyponym: search component. Meronym. sup.: spoken language recognition system. Def.: In the recognition component, intervals of the speech signal are mapped by probabilistic systems such as Hidden Markov Models, Neural Networks, Dynamic Programming algorithms, Fuzzy Logic knowledge bases, to word hypotheses; the resulting mapping is organised as a word lattice or word graph, i.e. a set of word hypotheses, each assigned in principle to a temporal interval in the speech signal. (Gibbon et al. 1997, p. 190)

recording studio
/rɪˈkɔːdɪŋ ˈstjuːdɪəʊ/, /rI'kO:dIN 'stju:dI@U/, [N: [N: recording][N: studio]], [plural: -s]. Domain: physical characterisation. Hyperonyms: recording room. Cohyponym: laboratory room, soundproof booth, anechoic chamber. Def.: Speech recordings may be made in a professional recording studio. The advantage of this type of recording environment is that it is widely available and that the recording location may be rented only for the recording sessions. The major disadvantage of using a recording studio is that the recording conditions and especially the acoustic conditions are not standardised in any way. Moreover, it will generally not be possible to design the acoustic environment of the recording room according to the needs of speech recordings. (Gibbon et al. 1997, p. 309/310)

redundancy
/rɪˈdʌndənsi/, /rI'dVnd@nsi/, [N: redundancy], [plural: none]. Domain: multimodal systems. Hyperonyms: cooperation type. Cohyponym: complementarity, equivalence, specialisation, concurrency, transfer. Def.: The same chunk of information is transmitted using more than one modality. E.g. A costumer saying "I want the second item on the right", simultaneously pointing in that direction..

Redundant Array of Inexpensive Disks
/ˈræpɪd əˈreɪ əv ɪnɪkˈspensɪv ˈdɪsks/, /'r{pId @'reI @v InIk'spensIv 'dIsks/, [N: [AJ: Redundant][N: Array][PREP: of][AJ: Inexpensive][N: Disks]], [plural: always plural]. Synonyms: RAID. Def.: In a RAID array, several hard disks are combined in such a way that failure or removal of a disk does not interrupt the operation of the array as a whole. This is possible by distributing data over the individual hard disks, and by data duplication. Several RAID levels have been specified. They differ in the degree of redundancy and safety.

register
/ˈredʒɪstə/, /'redZIst@/, [N: register], [plural: -s]. Hyperonyms: language variety. Meronym. sup.: natural language. Def.: An occupation-oriented functional language variety such as the language of .

registered speaker
/ˈredʒɪstəd ˈspiːkə/ , /'redZIst@d 'spi:k@/ , [N: [AJ: registered][N: speaker]], [plural: -s] . Domain: speaker recognition. Hyperonyms: speaker. Hyponyms: dependable speaker, mistaken speaker, violated speaker, casual registered speaker. Synonyms: reference speaker, valid speaker, authorised speaker, subscriber, client. Cohyponym: impostor, non-registered speaker. Def.: A speaker who belongs to the list of registered users for a given speaker recognition system (usually a speaker who is entitled to use the facilities, the access of which is restricted by the system). (Gibbon et al. 1997, p. 413)

Google Exhibit 1008
Page 0045

regular grammar
/ˈkɒntekst ˈfriː ˈgræmə/, /'kQntekst 'fri: 'gr{m@/, [N: [N: context][AJ: free][N: grammar]], [plural: -s]. Domain: language modelling. Hyperonyms: grammar, context-free grammar. Hyponyms: deterministic regular grammar, nondeterministic regular grammar. Cohyponym: finite-state automaton. Def.: A regular grammar is a set of rules which defines linear structures over strings of symbols from a vocabulary V. It is defined formally as a quadrulple <N,T,S,R>, where N is a finite set of nonterminal symbols in V, T is a finite set of terminal symbols in V, S is a start symbol (defining the root of the tree structures) in N, and R is a set of rules either of the form A -> gB or the form A -> Bg (but not mixed), where A is an element of N, g is a non-zero string of symbols from T. Regular grammars are also known as Type 3 grammars in the Chomsky hierarchy of formal grammars. Hidden Markov Models are stochastic (probabilistic) regular grammars in which terminal symbols and transitions are annotated with application probabilities on the basis of corpus analyses. Regular grammars are weakly equivalent to, and processed by, finite state automata. (Crystal 1988, p. 71)

rejection
/rɪˈdʒekʃən/, /rI'dZekS@n/, [N: rejection], [plural: -s]. Domain: speaker recognition. Hyperonyms: decision outcome (of a speaker recognition system). Hyponyms: false rejection. Cohyponym: acceptance. Def.: 1. Decision outcome which consists in refusing to assign a registered identity (or class) in the context of open-set speaker identification or classification, or which consists in responding negatively to a speaker (class) verification trial. 2. The decision of an automatic speech recognition (ASR) device that the input (or part of the input) cannot be mapped onto one or more words in the vocabulary with sufficient confidence. This results in the failure to recognise the input.

resistant speaker
/rɪˈzɪstənt ˈspiːkə/, /rI'zIst@nt 'spi:k@/, [N: [AJ: resistant][N: speaker]], [plural: -s]. Domain: speaker recognition. Hyperonyms: registered speaker. Synonyms: ram. Def.: A registered speaker with a low mistrust rate. (Gibbon et al. 1997, p. 433)

resonance disorder
/ˈrezənəns dɪsˈɔːdə/, /'rez@n@ns dIs'O:d@/, [N: [N: resonance][N: disorder]], [plural: -s]. Domain: corpora. Hyperonyms: speech disorder. Cohyponym: articulation disorder, voice disorder, language disorder, rhythm disorder. Def.: A resonance disorder involves lesions of the oral, nasal, or laryngeal cavities. Apart from functional causes, resonance disorders can result from surgical removal of the tonsils, a cleft palate, or nose polyps. (Gibbon et al. 1997, p. 114)

restricted language
/rɪˈstrɪktɪd ˈlæŋgwɪdʒ/, /rI'strIktId 'l{ngwIdZ/, [N: [AJ: restricted][N: language]], [plural: -s]. Hyperonyms: natural language. Def.: A variety of natural language which is restricted by externally imposed rules of use. These rules typically limit the vocabulary and the range of acceptable syntactic constructions. Restricted languages tend to be used in contexts where rapid, effective communication of a small set of basic facts is paramount, for example, in air traffic control. Because of the tightly constrained nature of restricted languages, they are seen by many to be good candidates for modelling interactive dialogue systems. However, this advantage must be weighed against the safety-critical function of many such languages in real use.

retroflex consonant
/ˈretrəʊfleks ˈkɒnsənənt/, /'retr@Ufleks 'kQns@n@nt/, [N: [AJ: retroflex][N: consonant]], [plural: -s]. Hyperonyms: consonant. Cohyponym: bilabial consonant, labiodental consonant, dental consonant, alveolar consonant, postalveolar consonant, palatal consonant, velar consonant, uvular consonant, pharyngeal consonant, glottal consonant. Def.: Retroflex consonant is a term used in the phonetic classification of consonant sounds on the basis of their place of articulation: it refers to a sound made when the tip of the tongue is curled back in direction of the front part of the hard palate. (Crystal 1988, p. 265)

Google Exhibit 1008
Page 0046

speaker labelling
/'spiːkə 'leɪbəlɪŋ/, /'spi:k@ 'leIb@lIN/, [N: [N: speaker][N: labelling]], [plural: -s]. Domain: speaker recognition. Hyperonyms: evaluation method, recognition task. Cohyponym: speaker matching, speaker alignment, speaker change detection. Def.: When the identity of the speakers taking part in a conversation is known, the goal is to localise when their successive interventions begin and end, including a possible outcome of 'none of the registered speakers', in case of open-set labelling. (Gibbon et al. 1997, p. 411)

speaker matching
/'spiːkə 'mætʃɪŋ/, /'spi:k@ 'm{tSIN/, [N: [N: speaker][N: matching]], [plural: -s]. Domain: speaker recognition. Hyperonyms: evaluation method, recognition task. Cohyponym: speaker labelling, speaker alignment, speaker change detection. Def.: The task of choosing a speaker in a closed-set of references which is most similar to a current speaker, even though it is known in advance that the applicant speaker is not a registered speaker. (Gibbon et al. 1997, p. 411)

speaker recognition
/'spiːkə rekəg'nɪʃən/, /'spi:k@ rek@g'nIS@n/, [N: [N: speaker][N: recognition]] , [plural: -s] . Domain: speaker recognition. Hyperonyms: evaluation method. Hyponyms: speaker verification, speaker identification, speaker classification; speaker class identification, speaker class verification; spoken language identification, spoken language verification.. Def.: Any decision-making process that uses some features of the speech signal to determine some information on the identity of the speaker of a given utterance.

speaker verification task
/'spiːkə verɪfɪ'keɪʃən 'tɑːsk/, /'spi:k@ verIfI'keIS@n 'tA:sk/, [N: [N: speaker][N: verification][N: task]], [plural: -s]. Domain: speaker recognition. Hyperonyms: recognition task. Cohyponym: speaker identification task. Def.: The speaker verification task is to decide whether or not the unlabelled voice belongs to a specific reference speaker. (Gibbon et al. 1997, p. 411)

speaker verification
/'spiːkə verɪfɪ'keɪʃən/ , /'spi:k@ verIfI'keIS@n/ , [N: [N: speaker][N: verification]] , [plural: -s] . Domain: speaker recognition. Hyperonyms: evaluation method; speaker recognition. Synonyms: speaker authentication. Cohyponym: speaker identification, speaker classification. Def.: Any decision-making process that uses some features of the speech signal to determine whether the speaker of a given utterance is a particular person, whose identity is specified.

specialisation
/speʃəlaɪ'zeɪʃən/, /speS@laI'zeIS@n/, [N: specialisation], [plural: none]. Domain: multimodal systems. Hyperonyms: cooperation type. Cohyponym: complementarity, redundancy, equivalence, concurrency, transfer. Def.: A specific chunk of information is always transmitted using the same modality. Specialisation may also manifest itself in user preferences, for example, if users consistently prefer speech over other input modalities for certain tasks. E.g. An information kiosk offers different services which are selected by touching the corresponding button..

specific lexicon theory
/spə'sɪfɪk 'leksɪkən 'θiːəri/, /sp@'sIfIk 'leksIk@n 'Ti:@ri/, [N: [AJ: specific][N: lexicon][N: theory]], [plural: y/-ies]. Domain: lexicon. Hyperonyms: lexicon theory. Cohyponym: general lexicon theory. Def.: A specific lexicon formulated in a lexicon formalism on the basis of a lexicon model.

Google Exhibit 1008
Page 0047

specification and description language
/spesɪfɪ'keɪʃən 'ænd dɪ'skrɪpʃən 'læŋgwɪdʒ/, `/spesIfI'keIS@n '{nd dI'skrIpS@n 'l{NgwIdZ/`, [N: [N: Specification][C: and][N: Description][N: Language]], [plural: -s]. Domain: interactive dialogue systems. Synonyms: SDL. Def.: A graphical language for describing state transition diagrams for event-driven systems. It was standardised by CCITT. (Gibbon et al. 1997, p. 573)

speech act
/'spiːtʃ 'ækt/, `/'spi:tS '{kt/`, [N: [N: speech][N: act]], [plural: -s]. Hyperonyms: action. Def.: A speech act is the informational action that a speaker effects by producing an utterance. For example, asking a question, offering information, and making a promise are three different types of speech act. The basic idea of speech acts is vitally important in work on dialogue systems. Speech acts serve as the base level of categorisation for dialogue work (in much the way that word classes have that function at the lexical level). So, for example, dialogue grammars can be written which describe well-formed sequences of speech acts. Many researchers working on interactive dialogue systems wish to use the notion of speech act without enlisting the whole philosophical apparatus of Speech Act Theory; for this purpose the term dialogue act has been coined and is steadily growing in acceptability.

speech aware consumer electronics
/'spiːtʃ ə'weə kən'sjuːmərɪlek'trɒnɪks/, `/'spi:tS @'we@ k@n'sju:m@rIlek'trQnIks/`, [N: [N: speech][AJ: aware][N: consumer][N: electronics]], [plural: none]. Domain: consumer off-the-shelf products. Def.: Mobile telephone, video recorders, TVs, car radios, etc. that can be controlled by speech input. At the time of writing, most of these systems are under development, and only a few are already available.

speech database
/'spiːtʃ 'deɪtəbeɪs/, `/'spi:tS 'deIt@beIs/`, [N: [N: speech][N: database]], [plural: -s]. Hyperonyms: database. Def.: A systematic database containing a collection of speech signals with header information, transcriptions and signal annotations.

speech output assessment
/'spiːtʃ 'aʊtpʊt ə'sesmənt/ , `/'spi:tS 'aUtpUt @'sesm@nt/` , [N: [N: speech][N: output][N: assessment]], [plural: -s] . Domain: speech synthesis. Hyperonyms: evaluation method. Synonyms: speech output testing. Def.: Determination of the quality of (some aspect(s) of) a speech output system.

speech output evaluation
/'spitʃ 'aʊtpʊt ɪvælju'eɪʃən/, `/'spitS 'aUtpUt Iv{lju'eIS@n/`, [N: [N: speech][N: output][N: evaluation]] , [plural: -s] . Domain: speech synthesis. Hyperonyms: evaluation method. Synonyms: speech output testing. Def.: Determination of the quality of (some aspect(s) of) a speech output system.

speech output system
/'spiːtʃ 'aʊtpʊt 'sɪstəm/, `/'spi:tS 'aUtpUt 'sIst@m/`, [N: [N: speech][N: output][N: system]], [plural: -s]. Domain: speech synthesis. Hyponyms: text-to-speech system, concept-to-speech system. Def.: A device, either a dedicated machine or a computer programme, that produces signals that are intended to be functionally equivalent to speech produced by humans. In the present state of affairs speech output systems generally produce audio signals only, but laboratory systems are being developed that supplement the audio signal with the visual image of the (artificial) talker's face.

speech output testing
/'spiːtʃ 'aʊtpʊt 'testɪŋ/ , `/'spi:tS 'aUtpUt 'testIN/` , [N: [N: speech][N: output][N: testing]] , [plural: -s] . Domain: speech synthesis. Hyperonyms: testing procedure. Def.: Determination of the quality of (some aspect(s) of) a speech output system.

Google Exhibit 1008
Page 0048

Spoken Language Processing
/ˈspəʊkən ˈlæŋgwɪdʒ ˈprəʊsesɪŋ/, /'sp@Uk@n 'l{NgwIdZ 'pr@UsesIN/, [N: [AJ: spoken][N: language][N: processing]], [plural: none]. Domain: processing. Hyperonyms: language processing. Hyponyms: automatic speech recognition, automatic speech synthesis, speaker recognition. Synonyms: SLP, speech processing. Cohyponym: natural language processing, text processing, word processing. Def.: Spoken language processing is an area of research and development in the field of human language technologies concerned with input and output systems which process natural human speech.

spoken language technology
/ˈspəʊkən ˈlæŋgwɪdʒ tekˈnɒlədʒi/, /'sp@Uk@n 'l{NgwIdZ tek'nQl@dZi/, [N: [AJ: spoken][N: language][N: technology]], [plural: y/-ies]. Hyperonyms: technology, human language technology. Hyponyms: speech input technology, speech output technology. Synonyms: spoken language technology, spoken language processing, SLP. Cohyponym: text technology, natural language processing, NLP. Def.: The discipline concerned with the research and development of spoken language input and output systems, using contributions from the neighbouring disciplines of acoustics, electrical engineering, statistics, phonetics, natural language processing, and involving system requirements specification, design, implementation and evaluation, corpus and linguistic resource processing, and consumer oriented product evaluation. (Gibbon et al. 1997, p. 578)

spoken language text corpus
/ˈspəʊkən ˈlæŋgwɪdʒ ˈtekst ˈkɔːpəs/, /'sp@Uk@n 'l{NgwIdZ 'tekst 'kO:p@s/, [N: [AJ: spoken][N: language][N: text][N: corpus]], [plural: spoken language text corpora]. Domain: corpora. Hyperonyms: corpus. Def.: A spoken language text corpus is a collection of data not taken from existing texts but from speech data that are written down in some orthographic or non-orthographic form in order to become part of a data collection. (Gibbon et al. 1997, p. 81)

spoken language verification
/ˈspəʊkən ˈlæŋgwɪdʒ verɪfɪˈkeɪʃən/ , /'sp@Uk@n 'l{NgwIdZ verIfI'keIS@n/ , [N: [AJ: spoken][N: language][N: verification]], [plural: -s]. Domain: speaker recognition. Hyperonyms: evaluation method; decision-making process. Cohyponym: spoken language identification. Def.: Any decision-making process that uses some features of the speech signal to determine whether the language spoken in a given utterance is a particular language.

spontaneous speech
/spɒnˈteɪnɪəs ˈspiːtʃ/, /spQn'teInI@s 'spi:tS/, [N: [AJ: spontaneous][N: speech]], [plural: none]. Domain: speech recognition, consumer off-the-shelf products. Hyperonyms: speaking style. Cohyponym: read speech, dictation speech. Def.: This is the most representative speaking style. In everyday life people communicate by talking spontaneously to each other. For a recognition system this talking style is very difficult to recognise. The style is characterised by level variations and the use of unpredictable intonation, hesitations, corrections, and incomplete sentences that are often grammatically incorrect.

SQL
/ˈeskjuːˈel/, /'eskju:'el/, [N: SQL], [plural: none]. Hyperonyms: query language. Synonyms: Standard Query Language. Cohyponym: OSQL. Def.: SQL is the de facto standard language for relational databases, and SQL-3 is currently being standardised by the ISO; important new features are the computation of transitive closure, and object-oriented concepts.

Standard Generalized Markup Language
/ˈstændəd ˈdʒenərəlaɪzd ˈmɑːkʌp ˈlæŋgwɪdʒ/, /'st{nd@d 'dZen@r@laIzd 'mA:kVp 'l{NgwIdZ/, [N: [AJ: Standard][AJ: Generalized][N: Markup][N: Language]], [plural: none]. Hyperonyms: formal language. Hyponyms: eXtended Markup Language (XML). Synonyms: SGML. Def.: An ISO standard for markup (annotation) which describes the structure of a text.

Google Exhibit 1008
Page 0049

standard procedure
/'stændəd/, /'st{nd@d/, [N: standard], [plural: -s]. Domain: . Cohyponym: ad hoc procedure. Def.: 1. A laboratory procedure conforming to agreed professional best practice. 2. A procedure defined by a standardisation organisation such as ISO, DIN, BSI.

Standard Query Language
/'stændəd 'kwiɪəri 'læŋgwɪdʒ/, /'st{nd@d 'kwi:@ri 'l{NgwIdZ/, [N: [AJ: Standard][N: Query][N: Language]], [plural: none]. Hyperonyms: query language. Synonyms: SQL. Cohyponym: OSQL. Def.: The de facto standard language for relational databases, and SQL-3 is currently being standardised by the ISO; important new features are the computation of the transitive closure, and object-oriented concepts.

static-dynamic representation
/'stætɪk 'daɪ'næmɪk reprɪzen'teɪʃən/, /'st{tIk 'daI'n{mIk reprIzen'teIS@n/, [N: [AJ: static][AJ: dynamic][N: representation]], [plural: -s]. Domain: multimodal systems. Hyperonyms: output modality representation. Cohyponym: linguistic representation, analogue representation, iconic representation, arbitrary representation. Def.: The representation is considered static when it can be perceived in an identical form for a certain time. If the representation changes continuously over time, it is called a dynamic representation. A blinking icon is considered static while a movie or music will be characterised as dynamic.

stem lexicon
/'stem 'leksɪkən/, /'stem 'leksIk@n/, [N: [N: stem][N: lexicon]], [plural: stem lexica, -s]. Domain: lexicon. Hyperonyms: lexicon. Cohyponym: morph lexicon, morpheme lexicon, fully inflected form lexicon. Def.: A lexicon in which the basic lexical key or lemma is the stem, which is represented in some kind of normalised notation (e.g. 'infinitive' for verbs, 'nominative singular' for nouns, in a standardised orthographic representation). (Gibbon et al. 1997, p. 199)

stem
/'stem/, /'stem/, [N: stem], [plural: -s]. Domain: lexicon. Cohyponym: affix. Def.: In the most general usage, a stem is any uninflected item, whether morphologically simple or complex. However, intermediate stages in word formation by affixation, and in the inflection of highly inflected languages, are also called stems. The smallest stem is a phonological lexical morph or an orthographic lexical morph, i.e. the phonological or orthographic realisation of a lexical morpheme. Stems may vary in different inflectional contexts. (Gibbon et al. 1997, p. 199)

stochastic grammar
/stə'kæstɪk 'græmə/ , /st@'k{stIk 'gr{m@/ , [N: [AJ: stochastic][N: grammar]], [plural: -s] . Domain: language modelling. Def.: A stochastic grammar is a stochastic language model that is based on a (context free) grammar; the grammar rules are assigned probabilities such that each word string generated by the grammar has a non-zero probability.

stochastic language model
/stə'kæstɪk 'læŋgwɪdʒ 'mɒdəl/, /st@'k{stIk 'l{NgwIdZ 'mQd@l/, [N: [AJ: stochastic][N: language][N: model]] , [plural: -s]. Domain: language modelling. Hyperonyms: language model. Cohyponym: grammar based language model. Def.: A stochastic language model is a language model that assigns probabilities to the allowed word sequences; typically all word sequences have a non-zero probability.

stochastic speech recognition system
/stə'kæstɪk 'spiːtʃ rekəg'nɪʃən 'sɪstəm/, /st@'k{stIk 'spi:tS rek@g'nIS@n 'sIst@m/, [N: [AJ: stochastic][N: speech][N: recognition][N: system]], [plural: -s]. Domain: language modelling, speech recognition. Hyperonyms: speech recognition system. Def.: A stochastic speech recognition system relies on stochastic models which are estimated or trained with (very) large amounts of speech, using some statistical optimalisation procedure. (Gibbon et al. 1997, p. 94) E.g. Hidden Markov Model (HMM), neural network.

Google Exhibit 1008
Page 0050

text-dependent speaker recognition system
/'tekst dɪ'pendənt 'spiːkə rekəg'nɪʃən 'sɪstəm/, **/'tekst dI'pend@nt 'spi:k@ rek@g'nIS@n 'sIst@m/**, [N: [N: text][AJ: dependent][N: speaker][N: recognition][N: system]], [plural: -s]. Domain: speaker recognition. Hyperonyms: speaker recognition system. Cohyponym: text-independent speaker recognition system . Def.: A speaker recognition system for which the training and test speech utterances are composed of exactly the same linguistic material, in the same order (typically, a password).

text-independent speaker recognition system
/'tekst ɪndɪ'pendənt 'spiːkə rekəg'nɪʃən 'sɪstəm/ , **/'tekst IndI'pend@nt 'spi:k@ rek@g'nIS@n 'sIst@m/** , [N: [N: text][AJ: independent][N: speaker][N: recognition][N: system]] , [plural: -s] . Domain: speaker recognition. Hyperonyms: speaker recognition system . Hyponyms: unrestricted text-independent speaker recognition system, event-dependent speaker recognition system. Cohyponym: text-dependent speaker recognition system. Def.: A speaker recognition system for which the linguistic content of test speech utterances varies across trials.

text-prompted speaker recognition system
/'tekst 'prɒmptɪd 'spiːkə rekəg'nɪʃən 'sɪstəm/ , **/'tekst 'prQmptId 'spi:k@ rek@g'nIS@n 'sIst@m/** , [N: [N: text][AJ: prompted][N: speaker][N: recognition][N: system]] , [plural: -s] . Domain: speaker recognition. Hyperonyms: speaker recognition system . Cohyponym: voice-prompted speaker recognition system, unprompted speaker recognition system. Def.: A speaker recognition system for which, during the test phase, a written text is prompted (through an appropriate device) to the user, who has to read it aloud.

text-to-speech system
/'tekst tə 'spiːtʃ 'sɪstəm/ , **/'tekst t@ 'spi:tS 'sIst@m/** , [N: [N: text][PREP: to][N: speech][N: system]], [plural: -s]. Domain: speech synthesis. Hyperonyms: speech output system. Cohyponym: concept-to-speech system. Def.: Speech output system that converts orthographic text (generally stored in a computer memory as ASCII codes) into speech.

text-to-visual-speech face synthesis
/'tekst tə 'vɪʒʊəl 'spiːtʃ 'feɪs 'sɪnθəsɪs/, **/'tekst t@ 'vIZU@l 'spi:tS 'feIs 'sInT@sIs/**, [N: [N: text][PREP: to][AJ: visual][N: speech][N: face][N: synthesis]], [plural: none]. Domain: multimodal systems. Hyperonyms: face synthesis. Cohyponym: puppeteer control face synthesis, performance-driven face synthsis, audio-driven face synthesis. Def.: The input of the system is plain text. The input text is first decomposed into its phonetic representation. Information about phonemes and their duration are automatically generated from the text. Formants and other speech parameters (frequency, pitch, pitch range and so on) are then computed. The text-to-visual-speech technique is suited when parametric facial models are used. Parameters defining facial animation are added to the set of speech parameters: lip shape, facial expressions, jaw rotation, etc. As a novel approach, speech synthesis systems have been extended to include facial parameters in their speech output parameters. The parallel computation of the auditory and visual parameters ensures a perfect synchronisation of the two channels, which is an advantage of such a technique. But different sampling rates of the speech synthesiser and of the animation system have to be reconciled. While the animation system uses 25-30 frames/sec, an acceptable audio system requires at least 50-60 frames/sec. To avoid temporal aliasing effect of the visual images, motion blur between successive frames can be used. Parameter values driving the facial model are blurred with their neighborhood (corresponding to the precedent and successive frames) parameters using a Gaussian filter. Text-to-visual-speech systems may be enhanced by adding markers describing intonation, speech rhythm, type of voice to the input text. Speech would be of better quality and such parameters could be used to get a more complex facial animation. For example, accents could be synchronised with raised eyebrows and head nods. Different facial models corresponding to different types of voice have also been explored.

theory-based evaluation
/ˈθiːəri ˈbeɪst ɪvæljʊˈeɪʃən/, /'Ti:@ri 'beIst Iv{ljU'eIS@n/, [N: [N: theory][AJ: based][N: evaluation]], [plural: -s]. Domain: multimodal systems. Hyperonyms: evaluation methodology. Cohyponym: expert-based evaluation, user-based evaluation. Def.: Theory-based evaluation involves a designer or evaluator who models task and user, based on the system specification. This ultimately generates quantitative values for interaction times, learnability or usability of the evaluated system. The evaluation involves neither a user-computer interaction nor a system prototype.

time-locked model
/ˈtaɪm ˈlɒkt ˈmɒdəl/, /'taIm 'lQkt 'mQd@l/, [N: [N: time][AJ: locked][N: model]], [plural: -s]. Domain: multimodal systems. Hyperonyms: coarticulation model. Cohyponym: look-ahead model, hybrid model, expansion model. Def.: The time-locked model is based on the principle that an event starts from an inherent time (a locked time). The protrusion influence due to a vowel appears at a given time before the vowel.

ToBI
/ˈtəʊbi/, /'t@Ubi/, [N: ToBI], [plural: none]. Hyponyms: E_ToBI, GlaToBI, J_ToBI,. Synonyms: Tone and Break Indices. Def.: A system of prosodic transcription which concentrates exclusively on represeting perceived pitch patterns (tones, 'To') in terms of target tone heights (usually two) and a hierarchy of boundary indices ('BI'). It has become perhaps the most popular and consistently applicable variety of prosodic transcription.

topic identification
/ˈtɒpɪk aɪdentɪfɪˈkeɪʃən/, /'tQpIk aIdentIfI'keIS@n/, [N: [N: topic][N: identification]], [plural: -s]. Hyperonyms: task. Synonyms: topic spotting. Def.: The determination of the topic of some speech or text material.

topline reference (condition)
/ˈtəplaɪn ˈrefərəns/, /'t@plaIn 'ref@r@ns/, [N: [N: topline][N: reference][N: condition]], [plural: -s]. Domain: speech synthesis. Hyperonyms: speech output. Cohyponym: baseline reference condition. Def.: Speech output that represents optimum performance, typically by a professional human talker.

tracking accuracy
/ˈtrækɪŋ ˈækjʊrəsi/, /'tr{kIN '{kjUr@si/, [N: [N: tracking][N: accuracy]], [plural: y/-ies]. Domain: multimodal systems. Def.: Percent deviation from true (facial) feature position.

tracking success
/ˈtrækɪŋ səkˈses/, /'tr{kIN s@k'ses/, [N: [N: tracking][N: success]], [plural: -es]. Domain: multimodal systems. Def.: Ratio of time when feature is tracked and time when feature is lost.

training
/ˈtreɪnɪŋ/, /'treInIN/, [N: training], [plural: -s]. Domain: speech recognition. Hyperonyms: process. Def.: The process in which a speech recognition system learns the pronunciation of words to be recognised at a later instance.

transaction
/trænsˈækʃən/, /tr{ns'{kS@n/, [N: transaction], [plural:s]. Domain: interactive dialogue systems. Meronym. sup.: dialogue . Def.: The part of a dialogue devoted to a single high-level task (for example, making a travel booking or checking a bank account balance). Atransaction may be coextensive with a dialogue, or a dialogue may consist of more than one transaction.

Google Exhibit 1008
Page 0052

TRANSCRIBER
/trænˈskraɪbə/, /tr{n'skraIb@/, [N: TRANSCRIBER], [plural: none]. Hyperonyms: annotation tool. Def.: A public domain tool for segmenting, labelling, and transcribing speech. It is written in Tcl/tk script language and is freely available as free software. TRANSCRIBER allows segmenting, labelling, and transcribing long duration signals. The output is in a standard SGML format. Multiple languages are supported. The tool can be ported to various platforms and is very flexible so that new functions can be easily added.

transcription
/trænˈskrɪpʃən/, /tr{n'skrIpS@n/, [N: transcription], [plural: -s]. Hyperonyms: notation. Hyponyms: phonetic transcription, phonemic transcription. Cohyponym: orthographic transcription. Def.: A method of writing down the names of speech sounds in a systematic and consistent way. (Crystal 1988, p. 313)

transfer
/ˈtrænsfɜɪ/, /'tr{nsf3:/, [N: transfer], [plural: none]. Domain: multimodal systems. Hyperonyms: cooperation type. Cohyponym: complementarity, redundancy, equivalence, specialisation, concurrency. Def.: A chunk of information in one modality triggers an event in another modality. E.g. In hypermedia interfaces: a mouse click provokes the display of an image..

trigram
/ˈtraɪgræm/, /'traIgr{m/, [N: trigram], [plural: -s]. Domain: language modelling. Cohyponym: zerogram, unigram, bigram. Def.: In language modelling a trigram is sequence of three words. (Gibbon et al. 1997, p. 94)

trill
/ˈtrɪl/, /'trIl/, [N: trill], [plural: -s]. Hyperonyms: consonant; manner of articulation. Cohyponym: plosive, nasal, tap, flap, fricative, lateral fricative, approximant, lateral approximant. Def.: Trill is a term used in the phonetic classification of consonant sounds on the basis of ther manner of articulation: it refers to any sound made by the rapid tapping of one organ of articulation against another. (Crystal 1988, p. 318) E.g. Example in English: [r].

turn duration
/ˈtɜɪn djuˈreɪʃən/, /'t3:n dju'reIS@n/, [N: [N: turn][N: duration]], [plural: –]. Domain: interactive dialogue systems. Def.: Turn duration is a measure of the average duration of one turn in a corpus of dialogues. (Gibbon et al. 1997, p. 606)

turn
/ˈtɜɪn/, /'t3:n/, [N: turn], [plural: -s]. Domain: interactive dialogue systems. Meronym. sup.: dialogue. Def.: A stretch of speech, spoken continuously by one party in a dialogue. A stretch of speech may contain several linguistic acts or actions. A dialogue consists of a sequence of turns produced alternately by each party. Turns are also known as utterances.

UCR
/ˈjuɪsiɪˈɑɪ/, /'ju:si:'A:/, [N: UCR], [plural: -s]. Hyperonyms: correction rate. Synonyms: user correction rate. Def.: Percentage of all user turns which are correction turns.

ultradirectional microphone
/ʌltrədaɪˈrekʃənəl ˈmaɪkrəfəʊn/, /Vltr@daI'rekS@n@l 'maIkr@f@Un/, [N: [AJ: ultradirectional][N: microphone]], [plural: -s]. Domain: physical characterisation. Hyperonyms: microphone. Cohyponym: unidirectional microphone, bidirectional microphone, omnidirectional microphone, pressure zone microphone, headset microphone. Def.: The ultradirectional microphone is designed for distant pickup, e.g. in film or TV productions. It strongly attenuates off-axis sound by means of multipath interference at a long slotted tube mounted in front of a unidirectional microphone. Compared to omni- and unidirectional microphones the sound quality is relatively poor since it has been traded against good directivity. The ultradirectional microphone is not recommended for high-quality speech recordings. (Gibbon et al. 1997, p. 305)

Google Exhibit 1008
Page 0053

unacquainted impostor
/ʌnə'kweɪntɪd ɪm'pɒstə/, /Vn@'kweIntId Im'pQst@/, [N: [AJ: unacquainted][N: impostor]], [plural: -s]. Domain: speaker recognition. Hyperonyms: intentional impostor. Cohyponym: acquainted impostor. Def.: An unacquainted impostor has never been in contact with the genuine or authentic user. (Gibbon et al. 1997, p. 422)

uncooperative speaker
/ʌnkəʊ'ɒpərətɪv 'spiːkə/, /Vnk@U'Qp@r@tIv 'spi:k@/, [N: [AJ: uncooperative][N: speaker]], [plural: -s]. Domain: speaker recognition. Hyperonyms: registered speaker. Cohyponym: cooperative speaker. Def.: When the user's goal and the system's purpose are inverse, an uncooperative (registered) speaker knows that he is being verified but wants the system to reject him. (Gibbon et al. 1997, p. 422)

unidirectional microphone
/juːnɪdaɪ'rekʃənəl 'maɪkrəfəʊn/, /ju:nIdaI'rekS@n@l 'maIkr@f@Un/, [N: [AJ: unidirectional][N: microphone]], [plural: -s]. Domain: physical characterisation. Hyperonyms: microphone. Hyponyms: cardioid microphone, hypercardioid microphone, supercardioid microphone. Cohyponym: bidirectional microphone, omnidirectional microphone, ultradirectional microphone, pressure zone microphone, headset microphone. Def.: The unidirectional type of microphone is most sensitive to sound arriving from one direction and more or less attenuates incident sound from other directions. Thus, unidirectional microphones will suppress intended sound when pointed at the wanted sound source, i.e. the speaker. (Gibbon et al. 1997, p. 303)

uniform language model
/'juːnɪfɔːm 'læŋgwɪdʒ 'mɒdəl/, /'ju:nIfO:m 'l{NgwIdZ 'mQd@l/, [N: [AJ: uniform][N: language][N: model]], [plural: -s]. Domain: language modelling. Hyperonyms: language model. Def.: Here, the idea is to use the same probability for all events; events can be either the words of the vocabulary or the sentences, if the number of sentences is limited. (Gibbon et al. 1997, p. 243)

Uniform Resource Locator
/'juːnɪfɔːm rɪ'sɔːs ləʊ'keɪtə/, /'ju:nIfO:m rI'sO:s l@U'keIt@/, [N: [AJ: Uniform][N: Resource][N: Locator]], [plural: -s]. Synonyms: URL. Def.: 1. Generally, a type of combined file name, data access pointer, and access parameters, for files located at arbitrary positions in a network. 2. Specifically, an Internet address, supplemented by detailed access information. A web client (WWW browser) requests a document via a URL from a WWW server. A URL has the form 'protocol://address:port/path/file#anchor?value_list'. 'protocol': an Internet protocol such as 'http', 'ftp', 'news', etc.; 'address': either an IP number or IP address; 'port': an operating system communications port number; 'path': a path name relative to the web server's root directory; 'file': a file name; 'anchor': a named position within the file; 'value_list': a list of attribute-value pairs written as 'attribute=value'. Attribute-value pairs are separated by '&'. URLs can be partial only; missing parts are substituted with default values by the server. The server interprets the URL and returns the requested document to the client.

unigram
/'juːnɪgræm/, /'ju:nIgr{m/, [N: unigram], [plural: -s]. Domain: language modelling. Cohyponym: bigram, trigram, zerogram. Def.: In language modelling: single word.

unimodal input event
/juːni'məʊdəl 'ɪnpʊt ɪ'vent/, /ju:ni'm@Ud@l 'InpUt I'vent/, [N: [AJ: unimodal][N: input][N: event]], [plural: -s]. Domain: multimodal systems. Hyperonyms: input event. Synonyms: monomodal input event. Cohyponym: multimodal input event. Def.: Set of user input events that belong together and are intended to convey a chunk of information, consisting of at least two parts in one modality.

Google Exhibit 1008
Page 0054

verb
/'vɜːb/, /'v3:b/, [N: verb], [plural: -s]. Domain: lexicon. Hyperonyms: lexical category. Cohyponym: noun, adjective, adverb. Def.: One of the four main lexical categories (parts of speech), typically occurring as main predicate of a sentence, modified by adverbs and by noun and sentence complements, and in many languages inflecting (inter alia) for person, number, and tense.

vernacular
/vəˈnækjʊlə/, /v@'n{kjUl@/, [N: vernacular], [plural: -s]. Hyperonyms: language variety. Cohyponym: standard, Lingua Franca (Crystal 1988). Meronym. sup.: natural language. Def.: Vernacular is a term used in sociolinguistics to refer to the indigenous language or dialect of a speech community. (Crystal 1988) E.g. vernacular of Liverpool, Berkshire, Jamaica (Crystal 1988).

violated speaker
/'vaɪəleɪtɪd 'spiːkə/, /'vaI@leItId 'spi:k@/, [N: [AJ: violated][N: speaker]], [plural: -s]. Domain: speaker recognition. Hyperonyms: registered speaker. Def.: The registered speaker owning the identity assigned erroneously to an impostor in open-set speaker identification. The registered speaker owning the identity claimed by a successful impostor in speaker verification. (Gibbon et al. 1997, p. 414)

viseme
/'vɪziːm/, /'vIzi:m/, [N: viseme], [plural: -s]. Domain: multimodal systems. Def.: Smallest unit of lip movement (while speaking) that can make a difference in meaning.

visual output device
/'vɪʒʊəl 'aʊtpʊt dɪ'vaɪs/, /'vIZU@l 'aUtpUt dI'vaIs/, [N: [AJ: visual][N: output][N: device]], [plural: -s]. Domain: multimodal systems. Hyperonyms: output device. Cohyponym: acoustic output device, haptic output device. Def.: Visual display by using a monitor is the most used means of communication via computer. Virtual reality, stereoscopic monitors, and immersive systems enhance spatial information by displaying data in 3D.

Viterbi alignment
/vɪ'tɜːbi əˈlaɪnmənt/, /vI't3:bi @'laInm@nt/, [N: [N: Viterbi][N: alignment]], [plural: -s]. Domain: language modelling. Hyperonyms: alignment algorithm. Synonyms: Viterbi decoding, Viterbi approximation, maximum approximation. Meronym. sup.: HMM recogniser. Def.: A widely used alignment algorithm that finds the best path through a probability graph. Viterbi alignment is usually applied to HMM output.

Viterbi approximation
/vɪ'tɜːbi əprɒksiˈmeɪʃən/, /vI't3:bi @prQksi:'meIS@n/, [N: [N: Viterbi][N: approximation]], [plural: -s]. Domain: language modelling. Hyperonyms: algorithm. Synonyms: maximum approximation, Viterbi alignment, Viterbi decoding. Def.: In the so-called Viterbi or maximum approximation, the sum over all paths is approximated by the path which has the maximum contribution to the sum.

Viterbi decoding
/vɪ'tɜːbi diːˈkəʊdɪŋ/, /vI't3:bi di:'k@UdIN/, [N: [N: Viterbi][N: decoding]], [plural: –]. Domain: language modelling. Hyperonyms: alignment algorithm. Synonyms: Viterbi alignment, Viterbi approximation, maximum approximation. Meronym. sup.: HMM recogniser. Def.: An algorithm used in pattern recognition that, under certain assumptions finds the best non-linear alignment of two sequences so as to maximise their similarity.

Google Exhibit 1008
Page 0055

vocabulary size
/vəˈkæbjʊləri ˈsaɪz/, /v@'k{bjUl@ri 'saIz/, [N: [N: vocabulary][N: size]], [plural: -s]. Domain: speech recognition, consumer off-the-shelf products. Hyponyms: active vocabulary size, passive vocabulary size, user vocabulary size, extension vocabulary size, exception vocabulary size. Synonyms: coverage, vocabulary coverage . Meronym. sub.: extensional coverage, intensional coverage. Def.: The size of the vocabulary is defined as the number of words that a speech recogniser can handle.

vocabulary
/vəˈkæbjʊləri/, /v@'k{bjUl@ri/, [N: vocabulary], [plural: y/-ies]. Hyponyms: active vocabulary, backup vocabulary. Synonyms: dictionary. Meronym. sup.: speech recogniser. Def.: The set of words in a lexicon, such as the set of words that an automatic speech recognition system is capable to recognise.

voice characteristic
/ˈvɔɪs kærəktəˈrɪstɪk/, /'vOIs k{r@kt@'rIstIk/, [N: [N: voice][N: characteristic]], [plural: -s]. Def.: Those aspects of speech which remain relatively constant over longer stretches of speech, and constitute the background against which segmental and prosodic variation is produced and perceived (e.g. mean pitch level, mean loudness, mean tempo, harshness, creak, whisper, tongue body orientation, dialect).

voice disorder
/ˈvɔɪs dɪsˈɔːdə/, /'vOIs dIs'O:d@/, [N: [N: voice][N: disorder]], [plural: -s]. Domain: corpora. Hyperonyms: speech disorder. Synonyms: disphonia. Cohyponym: articulation disorder, resonance disorder, language disorder, rhythm disorder. Def.: A voice disorder involves lesions of the vocal cords. The voice may emerge as a whisper (no vocal-cord vibration), for instance due to paralysis; or vocal-cord vibration may be present to some degree, but accompanied by excessive air flow (a "breathy" voice); or there may be irregular and therefore aperiodic vocal fold vibration, for instance due to the growth of abnormal tissue (nodules) on the vocal folds, resulting in a "hoarse" voice quality. Dysphonia may be caused by psychological and emotional factors, such as a severe shock, or by organic factors. A serious voice disorder is cancer of the vocal cords. (Gibbon et al. 1997, p. 115)

Voice Manager
/ˈvɔɪs ˈmænɪdʒə/, /'vOIs 'm{nIdZ@/, [N: [N: Voice][N: Manager]], [plural: -s]. Domain: consumer off-the-shelf products, speech recognition. Hyperonyms: tool. Meronym. sup.: speech recogniser. Def.: a commercial off-the-shelf (COTS) The voice commanding part of a recognition system that can track the contents of the currently active application (the window that has the focus), and dynamically adapts its active vocabulary. All normal Windows widgets can be read, such as buttons, menus, check boxes, radio buttons, pull down menus etc.

voice onset time
/vɔɪs ˈɒnset taɪm/, /vOIs 'Qnset taIm/, [N: [N: voice][N: onset][N: time]], [plural: -s]. Synonyms: voice onset time. Def.: The delay, which may be negative, between the release of a plosive and the onset of voicing; if there is perceptible delay, this may result in aspiration of the plosive, as with initial pre-vocalic plosives in English or German words.

voice-prompted speaker recognition system
/ˈvɔɪs ˈprɒmptɪd ˈspiːkə rekəgˈnɪʃən ˈsɪstəm/ , /'vOIs 'prQmptId 'spi:k@ rek@g'nIS@n 'sIst@m/ , [N: [N: voice][AJ: prompted][N: speaker][N: recognition][N: system]] , [plural: -s] . Domain: speaker recognition. Hyperonyms: speaker recognition system . Cohyponym: text-prompted speaker recognition system, unprompted speaker recognition system. Def.: A speaker recognition system for which, during the test phase, the user has to repeat a speech utterance, which he listens to through an audio device.

voicing decision
/ˈvɔɪsɪŋ dɪˈsɪʒən/, /'vOIsIN dI'sIZ@n/, [N: [N: voicing][N: decision]], [plural: -s]. Def.: The decision, used, for example, in certain speech coders, as to whether a particular section of speech is voiced or unvoiced.

Google Exhibit 1008
Page 0056

LIBRARY OF CONGRESS
0 007 111 679 7

HANDBOOK OF MULTIMODAL AND SPOKEN DIALOGUE SYSTEMS

Resources, Terminology and Product Evaluation

Dafydd Gibbon, Inge Mertins and Roger K. Moore (Eds.)

Dictation systems, read-aloud software for the blind, speech control of machinery, geographical information systems with speech input and output, and educational software with "talking head" artificial tutorial agents are already on the market. The field is expanding rapidly, and new methods and applications emerge almost daily. But good sources of systematic information have not kept pace with the body of information needed for development and evaluation of these systems. Much of this information is widely scattered through speech and acoustic engineering, linguistics, phonetics and experimental psychology.

THE HANDBOOK OF MULTIMODAL AND SPOKEN DIALOGUE SYSTEMS presents current and developing best practice in resource creation for speech input/output software and hardware. This volume brings experts in these fields together to give detailed "how to" information and recommendations on planning spoken dialogue systems, designing and evaluating audiovisual and multimodal systems, and evaluating consumer off-the-shelf products.

In addition to standard terminology in the field, the following topics are covered in depth:

- How to collect high quality data for designing, training and evaluating multimodal and speech dialogue systems.
- How to evaluate real life computer systems with speech input and output.
- How to describe and model human-computer dialogue precisely and in depth.

Also included:

- A fully searchable CD-ROM containing a hypertext version of the book in HTML format for fast look up of specific points, convenient desktop use, and lightweight mobile reference, and
- The first systematic medium scale compendium of terminology with definitions.

This handbook has been especially designed for the needs of development engineers, decision-makers, researchers and advanced level students in the fields of speech technology, multimodal interfaces, multimedia, computational linguistics, and phonetics.

SECS 565
0-7923-7904-7

ISBN 0-7923-7904-7
9 780792 379041