ARCHIVES OF ACOUSTICS
31, 4, 573–582 (2006)

C H R O N I C L E

Speech signal annotation, processing and synthesis
Poznań, Poland 12–13 September 2006

The conference *Speech signal annotation, processing and synthesis*, 11–13 September 2006 was organised by **Department of Phonetics of Adam Mickiewicz University** in Poznań, and **Institut für Kommunikationswissenschaften** in Bonn as a part of German-Polish project supported by **Alexander von Humboldt Stiftung**.

The conference was of interest to several scientific communities, including Speech Technologies, Linguistics, Phonetics, Audiology and Phoniatrics.

Specially, its scientific program covered a wide variety of topics related to the contemporary state of art on Speech Signal Annotation, Processing and Synthesis. While the meeting was focused mainly on the fields of speech annotation for needs of processing and synthesis of speech signals, contributions from many other areas of speech and language studies have also been accepted.

In order to support the multi-disciplinary goal of the conference, the international advisory committee included representatives of the different scientific communities. The conference was organised under auspices: *The International Speech Communication Association – ISCA* and *Polish Phonetic Association*

(http://www.staff.amu.edu.pl/~fonetyka/)

The scientific program included invited talks, presentations of papers in oral and poster sessions.

Main scientific areas:

- segmental and suprasegmental speech annotation,
- databases in speech synthesis and recognition,
- talking conditions,
- applications of medical phonetics: speech signal analysis based medical diagnostics.

Grażyna Demenko
President of the Polish Phonetic Association

Abstracts

1. Verification of a set of speech perception tests for children with a cochlear implant

BACHAN Jolanta, jolabachan@op.pl

Adam Mickiewicz University, Institute of Linguistics
Międzychodzka 5, 60-371 Poznań, Poland

This article concentrates on the results of the project work aiming at the verification of a set of speech perception tests for children with a cochlear implant created at the Adam Mickiewicz University in Poznań. The article outlines the key criteria for creating speech perception tests for children with normal hearing or children with hearing impairment and what childred demand from such tests. The key to the success is attractiveness, intelligibility and modelling of a dialogue between a child and a computer. Without these the child has to count on someone to help him (or her) understand the talking device.

⋆ ⋆ ⋆

2. Towards recognising speech gestures in discourse

CAMPBELL Nick, nick@nict.go.jp, nick@atr.jp

National Institute of Information and Communications Technology and
Advanced Telecommunications Research Institute
Spoken Language Communication Research Laboratory
International Keihanna Science City, Kyoto 619-0288, Japan

This paper describes how the lowest level of discourse information can be processed in a speech signal for the automatic annotation of discourse progress and for producing an estimate of speaker participation status. In a semi-formal round-table meeting situation, there is typically only one main speaker at any given moment, but several participants may be speaking simultaneously, expressing agreement (or otherwise), chatting, translating, etc., in addition to the main speaker. We are currently performing research into the technology to process this audio landscape, in order to detect the main speaker and to categorise the competing forms of speech in a given situation. Several speech gestures such as laughter, agreement and feedback-responses can be recognised, isolated and used to determine the progress of the meeting and the degrees and types of participation status among the members present. The technology exists to recognise these discourse events, but we still lack a model of their function in the mutual transfer of information through speech interaction.

⋆ ⋆ ⋆

3. Multilingual unit selection with the Boss system. Database annotation for Polish

DEMENKO Grażyna[(1)], lin@amu.edu.pl
Breuer Stefan[(2)]

[(1)]Adam Mickiewicz University, Institute of Linguistics
Międzychodzka 5, 60-371 Poznań, Poland
[(2)]Universität Bonn

The aim of the research concerns prosody annotation, especially for application in a corpus-based speech synthesis. In order to establish the rules of automatic intonation annotation and modelling, phonetically labeled speech database of 4 hours has been perceptually and acoustically analyzed. The speech material included different text types and prosodically rich phrases. The results of these analyses made it possible to define 6 intonation events to be used in modelling of the intonation contours for application in unit selection.

The naturally-sounding speech was produced by the corpus-based German speech synthesis system, BOSS. For prosody modeling, only fundamental types of structures were distinguished, such as word and

phrase accent placement or phrase boundary type. The results of the synthesis showed the outstanding quality of the database in phonetic-acoustic terms and confirmed the correctness of the automatic annotation on the segmental level and suprasegmental level.

⋆ ⋆ ⋆

4. Speech synthesis and language typology: insights from West African tone languages

GIBBON Dafydd, gibbon@uni-bielefeld.de

Universität Bielefeld
Fakultät für Linguistik und Literaturwissenschaft
Postfach 00131, D-33739 Bielefeld, Germany

Spoken Language systems, both for Automatic Speech Recognition and for Text To Speech Synthesis have been developed mainly for Indo-European languages, with a few exceptions such as Japanese and Chinese. There are notable exceptions, mainly in Africa, such as the prototype systems for isiZulu in South Africa, Gikuyo in Kenya, and Ibibio in Nigeria. There are obvious practical reasons for developing Spoken Language systems, particularly TTS systems, for disseminating information on agriculture, markets and health in pre-literate societies. But even more interesting in the long term is the theoretical interest in using TTS systems as an operational testbed for the kinds of linguistic theory which help to explain the potential of human cognition: operationalising a theory is a challenge in terms of truth and precision, which conventional linguistic and phonetic theory formation does not attain. Why is this an interesting question? The main reason is that the 6000 languages of the world are not constructed according to the same pattern as the Indo-European languages, or Japanese, or Chinese, and therefore systems for these languages do not generalise easily to other languages. Perhaps the most striking example of typological differences between languages is in lexical prosody. This typological difference will be the topic of the present contribution. To illustrate: many East Asian languages have lexical (i.e. phonemic) prosody, and African languages are also frequently mentioned as tone languages. However, tone functionality in African tone languages is fundamentally morphosyntactic rather than phonemic: (a) tonal pattern types are restricted to particular parts of speech, (b) tones may be inflectional and play a role in (c) derivational and (d) compounding word formation patterns, and (e) in syntactic phrasal templates. The aim of this paper is to document the morphosyntactic functionality of tones in African languages within a typological context as compared to East Asian tone languages such as Mandarin, and to develop finite state architectures for tone handling in practical Text-To-Speech synthesis in health and agriculture information projects in Ivory Coast and Nigeria. Morphosyntactic tone is illustrated for Ibibio (Lower Cross, South-Eastern Nigeria). The presentation will conclude with a discussion of experiences in the Ibibio TTS project, and an outline of the steps needed to go beyond "tone-deaf" speech synthesis in future projects.

⋆ ⋆ ⋆

5. Dynamic programming in post-processing of automatic segmentation of speech

SZYMAŃSKI Marcin, mszymanski@cs.put.poznan.pl
GROCHOLEWSKI Stefan

Poznań University of Technology, Institute of Computing Science
Piotrowo 2, 60-965 Poznań, Poland

Building unit-selection speech synthesisers requires large speech corpora, precisely annotated with phoneme boundary time-points. Manual segmentation of speech is a very laborious task, yielding the need for automatic segmentation algorithms. It was observed that the common HMM-based method is prone to systematical errors and some boundary refinement approaches, like boundary-specific correction, were introduced. Last year, a dynamic programming fine-tuning approach was proposed, that combined two sources information, boundary error distribution and boundary MFCC statistical models. In this paper we verify the usefulness of incorporating several other data, namely boundary energy dynamics models and the signal periodicity information.

⋆ ⋆ ⋆

6. Speech Efficiency Test as a tool in the diagnosis of extrapiramidal diseases. A case report of Multiple System Atrophy – Parkinson Type

GUSTAW Katarzyna[(1)], kasiagu@yahoo.ca
GONET Wiktor[(2)], gonet@klio.umcs.lublin.pl
[(1)]Maria Curie-Skłodowska University
Department of Logopedics and Applied Linguistics
Sowińskiego 17, Lublin, Poland and
IMW Lublin, Department of Neurodegenerative Disorders
[(2)]Maria Curie-Skłodowska University
Deptartment of English, Phonetics and Phonology Section
Pl. Marii Curie-Skłodowskiej 4, Lublin, Poland

The term 'multiple system atrophy' (MSA) describes a syndrome characterized by features overlapping with Shy-Drager syndrome, striatonigral degeneration, and olivopontocerebellar atrophy. MSA is characterized clinically by the combination of Parkinsonian, pyramidal, cerebellar, and autonomic symptoms. If Parkinsonism predominates in the clinical picture of the disease, the condition is named striatonigral degeneration or MSA-type P (P stands for Parkinsonism).

A 67 years old woman with a previous clinical diagnosis of striatonigral degeneration and a five year history of the symptoms had initially been misdiagnosed as suffering from Parkinson's disease. However, she did not respond well to levodopa and soon developed severe dysarthria. The neurological features, helpful in differentiating SND from other extrapiramidal disorders, included falling, dysarthria and dysphonia, respiratory stridor, hyperreflexia and ataxia. Cerebellar signs were well manifested, whereas autonomic symptoms were less severe.

Besides neurological and imaging study, a specially designed Speech Efficiency Test for dysarthria was performed. It revealed numerous speech deficiencies: a flat Fo contour, slow speech pace and a tendency to divide words into syllables. Vowels were centralized and reduced, obstruents articulated with impeded precision, and nasals produced with irregular soft palate timing. Glottal activity was characterized by lowered Fo and breathy phonation.

⋆ ⋆ ⋆

7. Talking blogs – an attempt to give weblogs a voice Adding TTS functionality to Wordpress

HELLMICH Arne, arne.hellmich@medi-ah.de

Bielefeld University, Department of Linguistics and Literary Studies
Postfach 100131, 33501 Bielefeld, Germany

In this paper I describe the development of a plug-in, which adds text-to-speech functionality to the blog publishing system Wordpress using synthetic voices developed by Cepstral, LLC. The RSS feed of Wordpress is used to generate audio files from the blog entries. Information on how to use the plug-in is provided and problems, which arise when converting blog entries into audio files using a speech synthesizer, are pointed out. Finally, a summary of the work is given and future developments are discussed.

⋆ ⋆ ⋆

8. Basic parameters in speech processing The need for evaluation

HÖGE Harald, harald.hoege@siemens.com

Siemens AG, Corporate Technology

As basic parameters in speech processing we regard pitch, duration, intensity, voice quality, signal to noise ratio, voice activity detection and strength of the Lombard effect. Taking into account also adverse conditions, the performance of many published algorithms to extract those parameters from the speech

signal automatically is not known. A framework based on competitive evaluation is proposed to push algorithmic research and to make progress comparable.

⋆ ⋆ ⋆

9. On intonation of question-type dialogue moves in Korean and Polish task-oriented dialogues Spontaneous speech analysis using perception modelling

Karpiński Maciej[(1)], maciej.karpinski@amu.edu.pl
Emilia Szalkowska[(2)], emilia.szalkowska@amu.edu.pl
[(1)]Adam Mickiewicz University, Institute of Linguistics
Międzychodzka 5, 60-371 Poznań, Poland
[(2)]Hankuk University of Foreign Studies, Department of Polish Studies, Korea

In this paper, a comparative study of the nuclear intonation in Korean and Polish questions in spontaneous speech is presented. Fluent realizations of three basic categories of question-type dialogue moves are extracted from the Pol'n'Asia Corpus of map task dialogues. In the analysis, Prosogram system by P. Mertens is employed for modelling pitch perception. Some details of pitch movement within prenuclear, nuclear and postnuclear syllables are described.

⋆ ⋆ ⋆

10. A study of chosen temporal relations within syllable structure in Polish

KLESSA Katarzyna, kasia@poczta.com.pl
ŚLEDZIŃSKI Daniel, danielsl@poczta.onet.pl
Adam Mickiewicz University, Institute of Linguistics
Międzychodzka 5, 60-371 Poznań, Poland

This paper presents data on the most common syllable patterns in Polish, based on a corpus of approximately 40 minutes of read speech and on a word list of nearly 700000 items. First, the results of statistical analysis concerning the frequency of occurrence of the possible syllable patterns in Polish are described. Then, selected problems connected with segmental duration within the syllable structure are addressed by presenting the results of duration measurements for particular elements of the syllable as related to stress and contextual conditions.

⋆ ⋆ ⋆

11. Disfluencies in Polish and Thai task oriented dialogue

KLEŚTA Janusz, janusz.klesta@amu.edu.pl
KARPIŃSKI Maciej, maciej.karpinski@amu.edu.pl
Adam Mickiewicz University, Institute of Linguistics
Międzychodzka 5, 60-371 Poznań, Poland

The present paper discusses the problem of addressing disfluencies contained in spontaneous speech corpora. Major types of disfluencies found in task-oriented dialogue recordings for two structurally different and geographically distant languages – Polish and Thai – are named, exemplified and tentatively categorized.

⋆ ⋆ ⋆

12. Text structure and prosody in Hungarian

KOUTNY Ilona
Adam Mickiewicz University, Institute of Linguistics
Międzychodzka 5, 60-371 Poznań, Poland

Frequent use of synthesized speech in different spheres of life necessitates a more natural sound. This means incorporation of the discourse level prosody, speaking styles and emotions. Text and discourse analyses can reveal some structural cues which have an influence on prosody for TtS systems. The current

investigations on written and spoken texts aims to determine the role of some discourse markers (conjunctions, pronouns) and question-answer pairs for discourse prosody.

⋆ ⋆ ⋆

13. Simple method of determining the voice similarity and stability by analyzing a set of very short sounds

ŁUKASZEWICZ Konrad

This paper presents a simple method of determining the voice similarity by analyzing a set of very short sounds. A large number of pitch-length sounds were extracted from natural voice signals from different realizations of open vowels 'a' and 'o'. The voice similarity was defined as the sum of single elementary similarities of short sound pairs. This method is oriented to the microphonemic speech synthesis based on waveform concatenation, and it could help to limit the time needed for database collection. This simple and low computational load speech synthesis method can be applied in small portable devices and used for the rehabilitation of speech-disabled people.

⋆ ⋆ ⋆

14. Will explicit criteria for boundary placement lead to higher inter-labeller agreement?

MACHAČ Pavel, SKARNITZL Radek, radek.skarnitzl@ff.cuni.cz, VOLÍN Jan

Charles University in Prague, Institute of Phonetics

The study investigates the effects of clear segmentation guidelines on inter-labeller agreement in boundary placement of Czech intervocalic plosives. With explicit criteria, the mean deviation in boundary placement was found to be only 1.55 ms, which is considerably lower than reported in literature. V-C boundaries invoked more discrepancies than the C-V ones. Boundaries which led to greatest deviations are discussed.

⋆ ⋆ ⋆

15. First steps in application of certain neural networks in phonetic phrase boundary recognition in Polish texts

MIGAS Arthur Z. P., arturmigas@polsl.pl

Silesian University of Technology, Institute of Computer Science
Akademicka 16, 44-100 Gliwice, Poland

The present paper presents author's continued research on automatic methods of phonetic phrase boundary recognition in Polish texts. The study concerns linguistic and automatic learning as well as the classification issues, since the solution has two stages: linguistic pre-processing followed by automatic learning and classification of objects. The research corpus is a set of newspaper texts with 10 K words (prepared by linguists). Gaps have been mapped to 4 depths of pauses (0...3). The corpus has been pre-processed (tagged) by LAS morphological analyser that uses up to 45 different POS symbols to substitute words. At the current stage, applicability of 2 kinds of neural networks as classification tool is researched. Conclusion is made that a 2-layer perceptron is better than radial network for this specific task. Neural networks are also compared to previously applied classifiers: AQ15 classifier and decision trees. The possibilities of improving the accuracy are discussed.

⋆ ⋆ ⋆

16. Exemplar-based speech representation

MÖBIUS Bernd

IMS, University of Stuttgart

Exemplar Theory posits that internal phonemic models emerge from storing in memory representations of large numbers of perceived acoustic realizations, or exemplars. There is evidence that these

exemplars and their phonetic details, rather than more abstract representations, are used in perception and production of process speech. In this talk I will sketch a computational model of the procedures that lead to the establishment of fully specified representations of speech in the framework of Exemplar Theory. The model is built on the hypothesis that exemplars of speech events are not concrete realizations but representations constructed through an internal analysis-by-synthesis process. This process starts from hypothesized lexical entries, takes landmark and context information as input, and results in fully specified exemplars of the pertinent linguistic unit.

⋆ ⋆ ⋆

17. Generalization in context sensitive grammars

PLUCIŃSKI Andrzej, apl@amu.edu.pl

Adam Mickiewicz University, Institute of Linguistics
Międzychodzka 5, 60-371 Poznań, Poland

An alternative for so far applied recognizing systems can be systems based on context-sensitive grammars. Precondition for success will be an implementation with great ability for generalization. We propose three methods of generalization based on the assumption that near stimuli invoke near or identical reactions.

⋆ ⋆ ⋆

18. Homophones in Polish

STEFFEN–BATÓG Maria[(1)]
BATÓG Tadeusz[(2)]

Adam Mickiewicz University
[(1)]Institute of Linguistics
Międzychodzka 5, 60-371 Poznań, Poland
[(2)]Faculty of Mathematics and Computer Sciences
Umultowska 87, 61-614 Poznań, Poland

This article introduces a more general concept of homophones than that used by linguists up until the present time. Four different kinds of homophones are defined, namely complete, contextual, facultative and facultative-contextual. A broad class of potential homophones that result from the increasing tendency to pronounce the word endings -ą, -ął, -oł identically is also discussed. The authors also announce their intention to complete a comprehensive dictionary of Polish homophones, which will have a wide range of applications.

⋆ ⋆ ⋆

19. Evaluation of the stylization of intonation contours

WAGNER Agnieszka, wagner@amu.edu.pl

Adam Mickiewicz University, Institute of Linguistics
Międzychodzka 5, 60-371 Poznań, Poland

This paper presents a new method of stylization of the intonation contours, which is the first step towards developing an intonation model for application in Polish unit selection speech synthesis. The paper starts with an overview of the existing F0 stylization algorithms underlying various intonation models, followed by an overview of descriptions of Polish intonation. Then the assumptions behind the new method and results of the quantitative and qualitative evaluation of the stylization accuracy are discussed. Finally, classification of pitch accents and boundary tones based on the parameterization is presented.

⋆ ⋆ ⋆

20. Visualisations of speech rhythm

WAGNER Petra

Universität Bonn

In recent years, several global rhythm measures have been proposed to correlate acoustic properties of the speech signal with the perceptual impression of listeners classifying languages as more or less "syllable timed" or "stress timed". While emphasizing the connection to rhythm related structural characteristics of the languages – such as phonotactic complexity, reduction and overall syllabic variability – they do not capture the fact that rhythm is characterised as a sequence of events in time. Another lack of global measures is the fact that rhythmic events are linked to different levels of the prosodic hierarchy (syllables, feet, phrases) and that the strength of these connections is language dependent.

A new visualisation of speech rhythm is presented, based on the durational relationships between subsequent syllables at different levels of the prosodic hierarchy. The two-dimensional visualisations show specific language differences which differentiate between stress-timed and syllable-timed languages. Furthermore, they show interesting differences between languages belonging to the same rhythm class: English shows a more distinct difference between adjoining stressed and unstressed syllables than German, which treats unstressed syllables more similar. Italian shows a very strong final lengthening effect but just as French, hardly differentiates between the neigbhoring stressed and unstressed syllables. Polish, which has been so far regarded as a rhythmically mixed language, is identified as syllable timed according to our classification. The visualisations furthermore help to identify rhythmic influences of a L1 on an L2.

⋆ ⋆ ⋆

21. Polish language dychotomic tests for speech audiometry
A study of people with normal hearing from various age groups

WISKIRSKA–WOŹNICA B.[(1)], WOJNOWSKI W.[(1)], OBRĘBOWSKI A[(1)]
PRUSZEWICZ A[(1)], DEMENKO G[(2)], WIKA T[(1)], KOMAR D.[(3)]

[(1)]Karol Marcinkowski Medical School in Poznań
Department and Hospital of Phoniatry and Audiology
[(2)]Adam Mickiewicz University
Department of Phonetics of the Linguistics Institute
[(3)]NZOZ MarkeMed in Poznań

Introduction: Evaluation of the objective results in the clinical examination of central auditory disorders requires the use of reliable language tests which provide the means for the estimation of patients' audio-verbal communicative skills.

Materials and methods: The authors present a new set of more difficult language tests in Polish, including a filtered speech test, numeral and verbal dichotic tests and a Calearo test. The tests were evaluated on a group of people with good hearing from various age groups.

Conclusions: The presented tests are the only existing ones created for the Polish language, which expand the diagnostic possibilities in the case of central auditory processing disorders.

It was found out that there is a right ear predominance in dichotic tests, which grows together with patient's age. This observation can be very beneficial in practice, allowing for hearing aids to be better fitted.

⋆ ⋆ ⋆

22. The acoustic voice and speech analysis in patients after partial tongue resection

WISKIRSKA–WOŹNICA B., ŚWIDZIŃSKI P., KRUK–ZAGAJEWSKA A.
ŻEBRYK–STOPA A., MIĘTKIEWSKA D.

University of Medical Sciences in Poznań
Chair and Department of Phoniatrics and Audiology

The physiological voice production depends on the condition of vibration apparatus in larynx as well as proper subglottic air pressure and volume of resonators as thorax, pharynx, oral and nasal cavity. There

is a possibility to change volume of this resonators and their mutual feedback due to proper movement of tongue, lips and mobility of mandibula. The shape and volume of the resonators has an effect on variability of the formants frequency, which can appear as articulation disorders and in this way can influence speech understanding.

The material of our investigation were 19 patients aged 46 to 67 years after partial tongue resection. All patients are divided into four groups depending on the type of surgery.

The method included perceptive voice and speech estimation (done by phoniatricians and speech therapists) and objective acoustic speech and voice analysis using Kay Elemetrics device and program. The character of formants F1 and F2 of the chosen consonants was presented as well as the type of changes in parameters describing voice in acoustic analysis in MDVProgram. This result of objective analysis was compared to subjective speech examination – articulation and speech understanding.

⋆ ⋆ ⋆

23. A case study on a virtual community: Forum of the Institute of Linguistics in Poznań

WŁODARCZYK Donata

The subject of this MA thesis is an existing virtual community of students of the Institute of Linguistics in Poznań. The aim of this research is to present the way this community functions and confront the online world of its users with its offline counterpart. The main emphasis will be laid on the users: their profile, the language they use and their observance of social norms of the group.

⋆ ⋆ ⋆

24. How to interpret a corpus

WIERZCHOŃ Piotr, DEMENKO Grażyna, lin@amu.edu.pl

Adam Mickiewicz University, Institute of Linguistics
Międzychodzka 5, 60-371 Poznań, Poland

The paper discusses all the stages involved in the preparation of electronic corpora to be used for the excerption of linguistic, especially lexical, information. The issues treated in detail include the input balancing, origin and selection criteria. Subsequently, techniques are described as "purification" of the raw, original electronic data for the most part downloaded from Internet resources. Another point explained in the paper is morphological analysis, or practical difficulties that emerged in the analysis of linguistic data contained in the corpus: (a) ambiguity of grammatical assignations; (b) inability to recognize correct spellings of words; (c) inability to recognize proper names; (d) inability to recognize abbreviations ending with a period; (e) incorrect analysis of words accompanied by punctuation marks, such as period, comma, exclamation mark, question mark, quotation marks etc.

The rules of phonetization, syllabification and accentuation in Polish have been formulated based on general linguistic theory and phonetic-acoustic analysis. All lexically accented syllables have been labelled according to LC_STAR language independent corpora specification.

⋆ ⋆ ⋆

25. K-means cluster analysis of melodic variation in Czech non-conclusive contours Comparison of the contemporary corpus data with the traditional descriptions of non-conclusive rises

VOLÍN Jan, jan.volin@ff.cuni.cz

Charles University in Prague, Institute of Phonetics
116 38 Praha 1, Czech Republic

An attempt to annotate prosodic features in the Prague corpus of read speech revealed that the typology of intonation contours provided in the phonetic literature did not reflect the tendencies in the use of specific

types of continuation rises in the material at hand. To map the situation we used repeated k-means cluster analyses of 252 cases under different conditions, with the aim to identify the possible candidate types of contours and to see how they were distributed. Among other things, the results confirm that k-means cluster analysis is a useful exploratory tool for the research tasks in which certain *a priori* knowledge and well-founded presumptions are available.

⋆ ⋆ ⋆

26. The use of statistics of Polish phonemes in speech recognition

ZIÓŁKO Bartosz[(1)], bziolko@cs.york.ac.uk, GAŁKA Jakub[(2)], jgalka@agh.edu.pl
MANANDHAR Suresh[(1)], suresh@cs.york.ac.uk, WILSON Richard C.[(1)], wilson@cs.york.ac.uk
ZIÓŁKO Mariusz[(2)], ziolko@agh.edu.pl

[(1)]University of York, Department of Computer Science
Heslington, YO10 5DD, York, UK
[(2)]AGH University of Science and Technology, Department of Electronics
Al. Mickiewicza 30, 30-059 Kraków, Poland

Statistical data on phonemes, useful in continuous speech recognition system, are presented. This paper explains the basics of a simple system for phonemes, diphones and triphones statistics estimation from a text corpus of Polish language. Obtained results are presented for exemplary text database. Possible application of statistics is suggested.

⋆ ⋆ ⋆