Proceedings
of the
ESCA Workshop on
Dialogue and Prosody

De Koningshof
Veldhoven, The Netherlands

September 1\textsuperscript{st} - 3\textsuperscript{rd}, 1999

Sponsored by:
European Speech Communication Association (ESCA)
VNC, Vlaams-Nederlands Comite voor Nederlandse Taal en Cultuur
Koninklijke Nederlandse Akademie van Wetenschappen
IPO, Center for Research on User-System Interaction
Eindhoven University of Technology (TUE)
Antwerp University (UIA)
The front cover shows a painting (1884) by Vincent van Gogh of "the Watermill at Coll" near Eindhoven; permission to reprint from the Lefrak Family Collection, New York, New York.
Organizing Committee

Marc Swerts  IPO, The Netherlands; UIA, Antwerpen, Belgium
Jacques Terken  IPO, The Netherlands
Bert van Coile  Lernout & Hauspie, Ieper, Belgium
Walter Daelemans  University of Brabant, Tilburg, The Netherlands; UIA, Antwerpen, Belgium
Werner Verhelst  Brussels Free University, Belgium

International Scientific Committee

Gösta Bruce  Lund University, Sweden
Nick Campbell  ATR, Kyoto, Japan
Carlos Gussengoven  Nijmegen University, The Netherlands
Julia Hirschberg  AT&T Research, Florham Park, USA
Daniel Hirst  University Aix-en-Province, France
Merle Horne  Lund University, Sweden
Jill House  University College London, UK
Stephen Isard  CSTR Edinburgh, UK
Joanna Moore  HCRC Edinburgh, UK
Anton Nijholt  Twente University, The Netherlands
Elmar Nöth  University Erlangen, Germany
Mari Ostendorf  Boston University, USA
Thomas Portele  Bonn University, Germany
Stephen Pulman  SRI Cambridge, UK
Angelien Sanderman  KPN Research, Leidschendam, The Netherlands
Elizabeth Shriberg  SRI Menlo Park, USA
David Traum  University of Maryland, USA
Bert van Coile  Lernout & Hauspie, Belgium
Gert Veldhuijzen van Zanten  IPO, Eindhoven, The Netherlands
Marilyn Walker  AT&T Reserch, Florham Park, USA
Introduction

Now that spoken dialogue systems are becoming more sophisticated, the utterances that may be exchanged between users and systems are beginning to become more complex and variable. In this context, increasing demands are placed on the way these systems deal with prosody, both with respect to generation and interpretation. If the prosody of the utterances generated by a system is appropriate, this will facilitate the comprehension of the system utterances by the user. Likewise, if a system exploits the prosodic information contained in the user utterances, it will be able to more adequately process the user utterances at different levels of representation.

The domain of dialogue modelling for spoken dialogue systems has received much attention in recent years, and in many cases reference has been made to the contribution that prosody might make to improving the performance of spoken dialogue systems. However, for experts in the area of prosody it is not always easy to link their work to the developments in the domain of dialogue modelling, whereas researchers working in the area of dialogue modelling often are rather naive with respect to prosodic modelling. We feel that bringing together researchers from both domains will provide either group with a better view of developments in the other domain. In our view this is a necessary pre-condition to enhance progress.

The ESCA Workshop on Dialogue and Prosody intends to provide a state-of-the-art overview of research in dialogue modelling and of attempts to improve the performance of spoken dialogue systems by means of the analysis and generation of prosodic features.

In line with the tradition of the ESCA Workshops, the workshop on Dialogue and Prosody has been organized in a (somewhat) remote location. De Koningshof in Veldhoven, The Netherlands, has been built in 1950 as a monastery by the Sisters of the Holy Heart, after their monastery in Moerdijk, Brabant had been destroyed in the Second World War. About 100 conventuals ran a secondary boarding school for about 300 girls, a very usual thing those days in the Catholic parts of Europe. In the sixties, changes within the congregation and in society at large led to the decline of the boarding school, and its purpose was changed into a congress center. Monasteries used to serve as a place of retreat, where people might withdraw for some time from the day-to-day worries and take a somewhat broader perspective on one's own efforts. The congress centers which constitute the modern analogon of these monasteries serve no other purpose. We trust that the relative seclusion provided by the Koningshof will encourage people to contemplate about the relations between their own research efforts and the broader developments in the field.

The topic areas of the workshop include intonational function, the synthesis of prosody in dialogue, the exploitation of prosody in speech recognition, and prosodic characteristics of backchanneling, barge-in and interruptions. Forty abstracts were submitted, each of which was reviewed by three members of the international scientific committee. After selection on the basis of recommendations by the international scientific committee and some additional withdrawals, twenty-nine papers remained. In order to ensure optimal information dissemination, most papers will be presented orally in plenary sessions, leaving ample space for discussion.
We thank ESCA for providing assistance in the organisation of the workshop. ESCA’s valuable guidelines for workshop organizers made it easy to get started and to go through the different stages in the preparations in due time. Also, we thank the other sponsors for providing the financial support that helped us to keep the registration fees at a minimum.

We thank the members of the International Scientific Committee for their willingness in reviewing abstracts and helping us to select the papers that constitute the raw materials for the workshop. We also thank the invited speakers for willing to contribute their expert views in the relevant domains. In a sense, they have become the victims of their scientific merits, which make them obvious targets for invitations to share their views with the profession, which they have accepted so benevolently.

Finally, we thank Mr. Edward Cortese of the Lefrak Organization for his kindness in granting us permission, on behalf of the Lefrak Family Collection, New York, New York, to reproduce the painting by Vincent van Gogh of "the Watermill at Coll" (1884), a historical building near Eindhoven.

Marc Swerts
Jacques Terken
Eindhoven, the Netherlands
September 1999
CONTENTS

Invited papers

Speaking in Time
H.H. Clark 1

Communication and prosody: Functional aspects of prosody
J. Hirschberg 7

Relating dialogue games to information state
S. Pulman 17

On the use of prosody in automatic dialogue understanding
E. Nöth, A. Batliner, V. Warnke, J. Haas, M. Boros, J. Buckow
R. Huber, F. Gallwitz, M. Nutt, H. Niemann 25

Session A: Intonation

Phonetic correlates of sentence type in Dutch: Statement, question
and command
V. van Heuven, J. Haan, R.S. Kirsner 35

Changing emotional tone in dialogue and its prosodic correlates
R. Cowie, E. Douglas-Cowie, A. Romano 41

Two dimensions of prominence
P. Wagner, T. Portele 47

Session B: Synthesis

A framework to allow dialogue systems to generate context-
sensitive prosody
P. Larrey 53

Speech rate control for dialogue speech synthesis based on the
prosodic structures
H. Kawanami, K. Hirose 59

The use of prosodic features to help users extract information from
structured elements in spoken dialogue systems
J. Hakulinen, M. Turunen, K.-J. Räähä 65

Session C: Discourse and Dialogue

Accentuation of domain-related information in Swedish dialogues
M. Horne, P. Hansson, G. Bruce, J. Frid, A. Jönsson 71

The meaning of melodic elements in Dutch
J. Caspers 77
Low-pitch regions as dialogue signals? Evidence from dialog-act and lexical correlates in natural conversation
_N. Ward_

**Session D: Poster session**

Syllable-based approach to automatic prosody detection: Applications for dialogue systems
_I. Kopecek_

Prosody and prompt design in a computer dialog system
_G. Ayers Elam, S.C. Wayland_

Prosodic correlates of discourse markers in dialogue
_P. Hansson_

On the relationship between the melodical structure and discourse functions of the particles NU and VOT in spontaneous Russian
_A. Kuosmanen_

The communicative effects of rising and falling pitch accents in British English and Dutch
_T. Rietveld, C. Gussenhoven, A. Wichmann, E. Grabe_

The role of prosodic cues in ASR, expert knowledge and human perception: A comparison of performance for French word recognition
_C. Lachaud, G. Caelen-Haumont, J. Pynte, R. Espesser_

Discourse effects on the prosodic properties of repetitions in human-computer interaction
_K. Fischer_

Prosody conveys information in severely impaired speech
_R. Patel_

Prosodic cues as basis for restructuring
_S. Jekat_

**Session E: Recognition**

Using high level dialogue information for dialogue act recognition using prosodic features
_H. Wright, M. Poesio, S. Isard_

Automatic intonation analysis using acoustic data
_K. Dusterhoff_
Using phrase accent information for dialog act recognition in
spontaneous German speech
M. Nutt, A. Batliner, V. Warnke, E. Nöth

Multi-lingual prosodic processing
J. Buckow, R. Huber, V. Warnke, A. Batliner, E. Nöth, H. Niemann

Prosodic information for integrated word-and-boundary recognition
F. Gallwitz, H. Niemann, E. Nöth, V. Warnke

Session F: Backchanneling, Barge-in, Interruptions

Prosodic correlates of disconfirmations

Measuring the effects of backchanneling in computerized oral
reading tutoring
G. Aist, J. Mostow

I said "TWO TI-CKETS": How to talk to a deaf wizard
H. Pirker, G. Loderer

An experimental study on the informational and grounding
functions of prosodic features of Japanese echoic responses
A. Shimojima, Y. Katagiri, H. Koiso, M. Swerts

Understanding recognition failures in spoken corrections in human-
computer dialogue
G.A. Levow