



TECHNICAL REPORT

Automatic speech recognition: Classification according to acoustic and linguistic indicators in real-life applications

INTERNATIONAL
ELECTROTECHNICAL
COMMISSION

ICS 33.160.01

ISBN 978-2-8327-0129-4

Warning! Make sure that you obtained this publication from an authorized distributor.

CONTENTS

FOREWORD.....	3
INTRODUCTION.....	5
1 Scope.....	6
2 Normative references	6
3 Terms and definitions	6
4 Use of automatic speech recognition (ASR) system.....	7
5 Needs for standards when using ASR.....	7
6 Definition and classification of factors affecting speech recognition	8
6.1 General.....	8
6.2 Acoustic indicators.....	8
6.2.1 Signal-to-noise ratio (SNR).....	8
6.2.2 Reflections	8
6.2.3 Reverberation.....	9
6.2.4 Data compression.....	9
6.3 Linguistic indicators	9
6.3.1 Unrestrained syntactical structure.....	9
6.3.2 Vocabulary list size.....	9
6.3.3 Homonyms	9
6.3.4 Multilingual words.....	9
6.3.5 Speaking speed.....	9
6.3.6 Accent	10
6.3.7 Speaking behavior.....	10
7 Existing International Standards	10
7.1 Inside IEC.....	10
7.2 In ISO/IEC Joint Technical Committee 1	10
8 Potential items in TC 100.....	10
Annex A (informative) A sample of an indicator set for classification.....	11
Bibliography.....	13
Table A.1 – Acoustic indicators and classification	11
Table A.2 – Linguistic indicators and classification.....	12

INTERNATIONAL ELECTROTECHNICAL COMMISSION

**AUTOMATIC SPEECH RECOGNITION:
CLASSIFICATION ACCORDING TO ACOUSTIC AND
LINGUISTIC INDICATORS IN REAL-LIFE APPLICATIONS**

FOREWORD

- 1) The International Electrotechnical Commission (IEC) is a worldwide organization for standardization comprising all national electrotechnical committees (IEC National Committees). The object of IEC is to promote international co-operation on all questions concerning standardization in the electrical and electronic fields. To this end and in addition to other activities, IEC publishes International Standards, Technical Specifications, Technical Reports, Publicly Available Specifications (PAS) and Guides (hereafter referred to as "IEC Publication(s)"). Their preparation is entrusted to technical committees; any IEC National Committee interested in the subject dealt with may participate in this preparatory work. International, governmental and non-governmental organizations liaising with the IEC also participate in this preparation. IEC collaborates closely with the International Organization for Standardization (ISO) in accordance with conditions determined by agreement between the two organizations.
- 2) The formal decisions or agreements of IEC on technical matters express, as nearly as possible, an international consensus of opinion on the relevant subjects since each technical committee has representation from all interested IEC National Committees.
- 3) IEC Publications have the form of recommendations for international use and are accepted by IEC National Committees in that sense. While all reasonable efforts are made to ensure that the technical content of IEC Publications is accurate, IEC cannot be held responsible for the way in which they are used or for any misinterpretation by any end user.
- 4) In order to promote international uniformity, IEC National Committees undertake to apply IEC Publications transparently to the maximum extent possible in their national and regional publications. Any divergence between any IEC Publication and the corresponding national or regional publication shall be clearly indicated in the latter.
- 5) IEC itself does not provide any attestation of conformity. Independent certification bodies provide conformity assessment services and, in some areas, access to IEC marks of conformity. IEC is not responsible for any services carried out by independent certification bodies.
- 6) All users should ensure that they have the latest edition of this publication.
- 7) No liability shall attach to IEC or its directors, employees, servants or agents including individual experts and members of its technical committees and IEC National Committees for any personal injury, property damage or other damage of any nature whatsoever, whether direct or indirect, or for costs (including legal fees) and expenses arising out of the publication, use of, or reliance upon, this IEC Publication or any other IEC Publications.
- 8) Attention is drawn to the Normative references cited in this publication. Use of the referenced publications is indispensable for the correct application of this publication.
- 9) IEC draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). IEC takes no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, IEC had not received notice of (a) patent(s), which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at <https://patents.iec.ch>. IEC shall not be held responsible for identifying any or all such patent rights.

IEC TR 63558 by IEC technical committee 100: Audio, video and multimedia systems and equipment. It is a Technical Report.

The text of this Technical Report is based on the following documents:

Draft	Report on voting
100/4214/DTR	100/4263/RVDTR

Full information on the voting for its approval can be found in the report on voting indicated in the above table.

The language used for the development of this Technical Report is English.

This document was drafted in accordance with ISO/IEC Directives, Part 2, and developed in accordance with ISO/IEC Directives, Part 1 and ISO/IEC Directives, IEC Supplement, available at www.iec.ch/members_experts/refdocs. The main document types developed by IEC are described in greater detail at www.iec.ch/publications.

The committee has decided that the contents of this document will remain unchanged until the stability date indicated on the IEC website under webstore.iec.ch in the data related to the specific document. At this date, the document will be

- reconfirmed,
- withdrawn, or
- revised.

INTRODUCTION

With the development of network and information technology, people are relying more and more on smart equipment, such as smart speakers, smart service robots and so on. Speech recognition technology is the main means to realize man-machine communication. Speech recognition is the process of converting a voice into digital data. Popular use in recent years has pushed improvements in its algorithm and increased accuracy. But the performance of different speech recognition solutions differs greatly and is sometimes a source of confusion for users. The factors used to evaluate the performance of speech recognition technology need more discussion.

This document mainly aims to set up a set of parameters which can be used to reflect the complexity of real-life applications, by means of a classification using scenarios.

AUTOMATIC SPEECH RECOGNITION: CLASSIFICATION ACCORDING TO ACOUSTIC AND LINGUISTIC INDICATORS IN REAL-LIFE APPLICATIONS

1 Scope

This document describes the factors related to classification of the real-life environment according to acoustic indicators and linguistic indicators. The set of factors can be used to describe complexities of use scenarios, from level 1 to 4, and can be helpful when setting up the testing environment.

This document applies for evaluating automatic speech recognition technology which is widely used for smart equipment, such as smart speakers.

2 Normative references

There are no normative references in this document.