FINAL
DRAFT

INTERNATIONAL
STANDARD

**ISO/FDIS
24620-3**

ISO/TC **37**/SC **4**

Secretariat: KATS

Voting begins on:
2020-11-24

Voting terminates on:
2021-01-19

Language resource management — Controlled human communication (CHC) —

Part 3:
Basic principles and methodology for controlled oral communication (COralCom)

Gestion des ressources linguistiques — Communication humaine contrôlée (CHC)

Partie 3: Principes de base et méthodologie de la communication orale contrôlée (COralCom)

Reference number
ISO/FDIS 24620-3:2020(E)

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT), see www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 37, *Language and terminology*, Subcommittee SC 4, *Language resource management*.

A list of all parts in the ISO 24620 series can be found on the ISO website.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html.

Introduction

No human is expected to master and speak several languages with the same level of competence as their own mother tongues. Faced with rapidly increasing multilingual situations, misunderstandings and confusion arise in daily communications, very often causing accidents and casualties. Industry and other business sectors, as well as different domains dealing with safety critical applications such as emergency services (police, fire, ambulance, maritime, etc.) require a precise and concise language to supplement the common use of natural human languages. Uncontrolled, these languages would allow non-parallel and non-comparable grammatical constructions, possessing inherent semantic ambiguities of variously phonetic, phonological, morphological, lexical or syntactic types.

Human communication is primarily oral, making use of the spoken medium of language.

Oral communication between native speakers, or non-native speakers, or a native speaker and a non-native speaker, can be disturbed due to different phenomena such as phoneme confusion and the resulting ambiguities due, for example, to multilingual communication or stressful situations. For an early study focussing on perceptual confusion among some consonants, see Reference [11]; for a more recent study concerning sources of variability in consonant perception, see Reference [12]. For a study and system on interferences in the domain of aeronautics and others, see References [6] and [7].

However, whatever the disturbance, the impaired communication results in confusion between distinctive phonetic features, phonemes and words including the problem of co-articulation (see Reference [24]). For early studies in the domain of phoneme confusion and interferences, and archiphonemes (see Reference [25]), norms, variants and common properties of all variants, distribution of variants as a norm as well as distinctive features of a phoneme, are presented and detailed, see References [13], [14] and [15].

This document deals with confusion and variable pronunciation or interpretation of distinctive phonetic features (including accents, stresses, and tones or lengths) and phonemes, but also homophones, homographs, and quasi-homophones and quasi-homographs.

The basic acoustic processes involved in speech production are the generation of sound sources and the filtering of these sources by the vocal tract. The sources are generated by the modulation of airflow through narrow constrictions produced at the larynx or in the vocal tract above the larynx, airflow that is usually the result of action of the respiratory system. The filtering of the sources is controlled by manipulating the shape of the vocal tract airway above the larynx, including the opening between the oral portion of the vocal tract and the nasal cavity. The acoustic properties of the sources and of the filtering of these sources can be varied by manipulating a number of different articulatory parameters relating to the laryngeal configuration, the positions and shapes of the lips, the tongue blade, the tongue body and other structures, and the stiffness of particular structures. Some of these manipulations lead to relative stable acoustic properties that are perceptually distinctive. These particular manipulations play a role in the selection of inventories of sounds consisting of discrete units such as features, segments, and words in terms of which language is structured (see Reference [17]).

Sounds of languages are described as sets of phonemes (see Reference [26]). All phonemes can be distinguished by at least one (acoustic/articulatory) feature. Every language takes a limited number of articulatory/acoustic features out of a virtually unlimited number of possibilities. For most known languages, the inventory ranges between thirteen and seventy-five phonemes (see Reference [19]). The phonetic characteristics of individual members of the inventory are, as a rule, given through matrices showing articulatory/acoustic features. A universal phonemic inventory has been provided (see References [20] and [22]). A phoneme system is the overall pattern of characteristics and relationships of the phonemes in the phonemic inventory of a given language. The phonological characteristics of the phonemes and their allophones are described by articulatory/acoustic features, the interrelationships between phonemes through oppositions. For a discussion of the importance of distinguishing between acoustics and articulatory features, a formant chart of the vowels of standard French and an illustration of the neutrality of distinctive features between production and perception, see Reference [21].

Human beings recognise in their respective language a phoneme and its variants which allows them to understand a word. In phonetics, sound variants belonging to one and the same phoneme are

called allophones, based on their distribution and phonetic similarity. A phonetic relationship in complementary distribution is, as a rule, one criterion for considering two sound variants as belonging to one and the same phoneme.

Considering that there are norms in languages, or that a language is a system of norms, of audible signs formed by the human organs and serving the purpose of communication, these norms can fulfil their task of serving communication only if both speaker and hearer relate to them within the same speech community. They are valid for the formation as well as for the perception of those signs.

The phonological system of a language is like a sieve through which everything that is said passes. Each person acquires the system of their mother tongue(s). When they hear another language spoken, they intuitively use the familiar 'phonological sieve' of their mother tongue(s) to analyse what has been said. Such filter is not suited for the foreign language, as typically it causes frequent mistakes and misinterpretations. The sounds of the foreign language receive an incorrect phonological interpretation since they are filtered through the respective 'phonological sieve' of one's own mother tongue(s) (see Reference [18]). Speakers and hearers are concerned by this 'phonological sieve'.

Language resource management — Controlled human communication (CHC) —

Part 3: Basic principles and methodology for controlled oral communication (COralCom)

1 Scope

This document provides basic principles and a methodology for establishing a specification for designing and constructing a formally defined, or controlled, system of oral communication that avoids or filters out phonetic interferences and confusions between words of the same language and between languages. The system is both abstracted from, and contextually situated in, the domains of industry, business or other technologies.

This document deals only with oral communication between native speakers, or non-native speakers, or a native speaker and a non-native speaker, who can be disturbed due to different phenomena, such as phoneme confusion, phonetic interferences and confusions between words (for example: homophony, quasi-homophony or co-articulation of the same language and/or different languages and the resulting ambiguities due, for example, to multilingual communication or stressful situations. This document deals with speakers and listeners without speech or hearing impediments[16], and does not include sign languages which have a phonological system equivalent to the system of sounds in spoken languages[23].

Foreseen applications are essentially in safety critical applications using human oral communication. This document is also applicable to other domains involving, for example, training and evaluation procedures and robots.

2 Normative references

There are no normative references in this document.

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

ISO and IEC maintain terminological databases for use in standardization at the following addresses:

— ISO Online browsing platform: available at https://www.iso.org/obp

— IEC Electropedia: available at http://www.electropedia.org/

3.1
controlled language
CL
system that limits language, written or spoken, to a set number of core lexical items and, usually, to a set of writing guidelines for grammar and style

Note 1 to entry: See also controlled (natural) language in ISO/TS 24620-1:2015.

3.2
distinctive feature
class of phonetically defined components of phonemes that function to distinguish meaning

Note 1 to entry: In contrast to redundant features, distinctive features constitute relevant phonological features.

Note 2 to entry: See also Reference [4], p. 134.

3.3
assimilation
articulatory adaptation of one sound to a nearby sound within a word or at the junction between words with regard to one or more features

Note 1 to entry: See also Reference [4], p. 40.

3.4
phoneme
smallest sound unit that can be segmented and that can function as a semantically distinctive unit

[SOURCE: Routledge Dictionary of Language and Linguistics, 1996]

3.5
homophone
one of two or more words that are pronounced the same but differ in meaning, origin, and sometimes spelling

[SOURCE: ISO 19104:2016, 4.15, modified — The Note 1 to entry has been removed.]

3.6
phoneme confusion
confusion due to a *phoneme* (3.4) approximately or incorrectly pronounced, and interpreted as another phoneme according to the mother tongue of the receptor

Note 1 to entry: Phonemes can exist in one language and not in other languages.

Note 2 to entry: Phonemes can be pronounced (spoken, emitted) with multiple accentuations, and be perceived by receptors (listeners) not necessarily receptive to the same phonetic and phonological systems.

3.7
quasi-homophones
words which differ by one or two *phonemes* (3.4)

Note 1 to entry: There can be one *phoneme* (3.4) more or less in one of the two quasi-homophones (e.g.: *aft-after*), one different phoneme (e.g.: *check-deck*, *feed-feet*), or 2 different phonemes (e.g.: *flap-slat*).

3.8
paronym
word for which the writing or pronunciation is very close to another word, but which has a different lexical meaning

3.9
interference
influence of one linguistic system on another in either the individual speaker or the speech community

Note 1 to entry: See also Reference [4], p. 235.

3.10
synonym
one of a set of different terms that refer to the same entity

[SOURCE: ISO/IEC 2382:2015, 2121523, modified — The Notes to entry have been removed.]

4 Motivation for controlled human communication

Globalisation, with the suppression of national borders and the pervasion of modern technology, necessitates that languages in the widest sense (including dialects) cohabit in everyday usage. Certain domains and especially those which are safety-critical need to block interferences between languages. Consider, for example, the situation where an air pilot listens to the co-pilot who reads out loud to him/her instructions which enable him/her to react to a problem during a flight. If the pilot's mother tongue is different from the co-pilot's, and as well as this, the messages are in American English, one can well appreciate the complexity of the problem. Confusions due to misunderstanding can at first sight seem anodyne; as history shows however, they have been at the origin of serious accidents.

The 1977 Tenerife air catastrophe is a convincing example of interference. The Dutch pilot of an aircraft, in declaring “we are at take-off”, because he made reference to a similar structure in Dutch he believed that he was indicating that he was about to take off, while in fact he was only indicating that he was ready to take off. He thus then incorrectly interpreted the control tower air traffic controller's approval as an authorisation for take-off. 583 persons lost their lives.

A few years later in 1983, after the crash of an aircraft, the black box revealed the Chinese co-pilot's last words: “What does 'pull-up' mean?”

The two following examples of phonetic confusion also show the need for controlled oral communication:

1) from ATC (a) to pilot (b):

   a) “Pass to the *left* of the tower”. (b) “Pass to the *west* of the tower”.

   The confusion was due to the fact that *west* and *left* meant two different sides of the tower.

2) to a Thai pilot:

   “descend to flight level two nine zero”. The pilot read back: “descend to flight level two five zero”.

   In this situation, the controlled language indicates that the digit *niner* replaces *nine*, which was not done and led to a confusion with *five.*

To avoid these sorts of problems, industry uses controlled languages.

Different domains are interested in controlled languages such as aircraft, meteorology, emergency services (police, fire, ambulance, maritime, etc.), etc. (see Figure 1 and Annex D, for more information see References [1] and [27]).

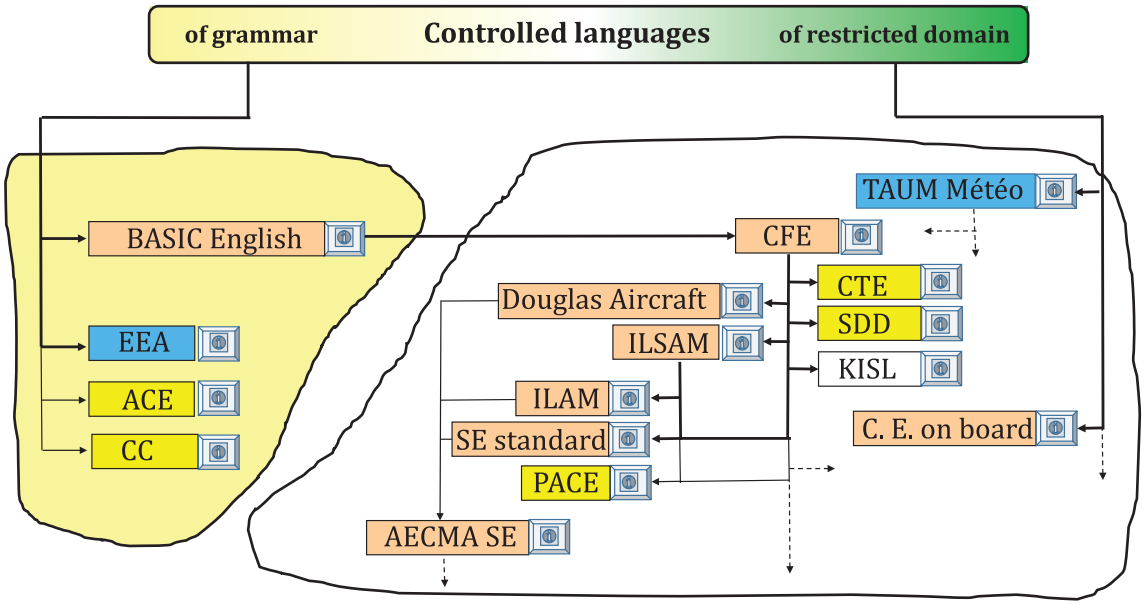

Figure 1 — Controlled languages

Controlled languages (CLs) are of vital interest (for safety and economic reasons, etc.) for industry. They have been created to resolve:

— problems of readability:

CLs reduce the complexity of syntactic structures of a text in order to increase its readability,

— problems of comprehensibility:

CLs do not accept lexical ambiguities; they thus increase the comprehensibility of a text, and

— problems of translatability:

their syntactic and semantic control facilitates the shift between two languages.

A controlled language attempts to reduce ambiguities, colloquialisms and synonyms (see Annex D and Reference [2]).

5 Basic principles and methodology

5.1 General

This clause provides the basic principles and the methodology for the overall specification aimed at the phonetic control of oral communication.

1) For the basic principles (see 5.3), various linguistic phenomena shall be used, especially concerning word formation and pronunciation that are represented with their distinctive phonetic features and phonological properties. Such use can be either context-free (abstracted from context) or context-sensitive (situated in a specific context) without relying on an existing lexicon which can neither be exhaustive nor be representative of all possible linguistic contexts.

2) The methodology (see 5.4) specifies a system of constraint rules that should/shall apply in a sequential order at each of the levels of linguistic analysis which have an impact on the generation and/or interpretation of sounds by a human being, such as phonetic and phonological, word-segmentation, morpho-syntactic, syntactic, semantic, and pragmatic.

These basic principles and the methodology are specified to provide an explanatory power to show how each of the linguistic components (minimal language sound units or distinctive features) or their combinations and transformations abstracted from, or embedded in, context shall be either used or avoided, so that no ambiguities of either intra- or inter-language are produced. In consequence, the methodology proposed uses the linguistic phenomena conformant to the basic principles for the production and generation of sounds and other relevant linguistic components, this instead of a lexicon, and specifies a system of constraint rules completed with an algorithm which, when applied, results in aiding avoiding ambiguities and confusion.

5.2 Problem: specific issues

The methodology for controlled oral communication focuses on the sound aspect of human communication. The basic problem is that oral messages interchanged between humans are pronounced with multiple accentuations and listened to by ears not necessarily receptive to the same phonetic system.

This problem is threefold as it concerns:

a) distinctive phonetic features and their contextually sensitive assimilatory variations,

b) phonemes that can be specific to each particular language such as, for example, Arabic, Chinese, English, French, Korean, etc., and

c) homophones and quasi-homophones engendering ambiguities.

The distinctive feature shall be used instead of listing all the possible words which cause confusion. A list of words is never exhaustive and is only relevant for a certain domain of application, and furthermore each word is dealt together with its context (environment in the sentence).

For this reason, this methodology is rule-based: instead of a lexicon, it establishes a system of ordered constraint rules based on linguistic phenomena that conform to the basic principles that shall be followed in order to choose the right word according to its context and to the specific application domain.

5.3 Principles

5.3.1 Overview

The principles addressing the three specific issues listed in 5.2 formulate rules for building a controlled oral lexicon. The process of this formulation starts with the analysis and description of distinctive (phonetic) features (5.3.2) that serve to build phonemes (5.3.3) and terminates with the formation of words (5.3.4) that include problematic homophones and quasi-homophones.

5.3.2 Distinctive features

Each of the sound segments such as stops [p], [t], and [k] or fricatives [f] and [v] is analysed as 'sets' of phonetic features, associated with their acoustic/articulatory manners or positions. The phonetic features of [p] and [v] are represented as below:

EXAMPLE 1 [p]: [+bilabial, -continuant, -voiced]

[v]: [+labiodental, +continuant, +voiced]

These two differ from each other in their position and manner of articulation. On the other hand, [p] and [b] differ only in their manner of articulation.

Given the physical state of the vocal tract, it is possible to predict the acoustic output using deterministic equations, which can be considered as mappings from articulation to acoustic output. The examples show the position in the vocal tract of some of its components (such as tongue, lips, vocal cords, ...), which play a role in distinguishing, for example, between [p] and [b].

EXAMPLE 2 [p]: [+bilabial, -continuant, -voiced]

[b]: [+bilabial, -continuant, +voiced]

EXAMPLE 3 [p]-[b] or [t]-[d] (-voiced vs. +voiced)

[p]-[t] (+bilabial vs +apico-alveolar)

5.3.3 From distinctive features to phonemes

Speakers of different languages can fail to perceive some of the distinctive features as being contrastive. Voicing, for instance, is such a case for speakers of Korean in which voicing is not a contrasting phonetic feature. So, for Koreans "pus" referring to a thick, yellowish liquid and "bus" referring to a vehicle are recognized as the same.

There is a vowel [y] which is analysed as having *closed, rounded*, and *anterior* features in contrast with the vowel [u], which is *closed, rounded*, and *non-anterior (back)*. This feature +/- anterior or posterior vs. anterior is a contrastive feature in French, thus differentiating the pronunciations of "dessus" (above) with [y] and "dessous" (below) with [u], while speakers of English fail to perceive the difference. Speakers of French recognize these two as two distinct sound segments, while speakers of English or other languages which have no [y] phoneme fail to recognize this vowel distinct from [u]. The word "bus" also has a different pronunciation in English and in French, [bʌs] with an *unrounded, neutral* and *central* vowel vs [bys].

The proximity between phonemes in the same language and in comparison with other languages is calculated (see for example Annex C).

Suprasegmental features such as accents, stresses, and tones or lengths can be distinctive and contrastive.

Examples where only the tone changes:

In Chinese:

Mā: 妈 (mom (US))

Má: 麻 (linen), 蔴 (hemp)

Mǎ: 马 (horse), 码 (code)

Mà: 骂 (curse)

In Swedish:

There are two tones, acute and grave, or tone 1 and tone 2. *buren* said with tone 1 means *cage*, with tone 2 means *carried*. Tone 2 is a compound tone. It falls from mid to low on the stressed syllable and then from high to mid on the following unstressed syllable. The classic minimal pair demonstrating the phonemic status of the tones in Swedish is *anden*. Said with tone 1, it means *duck*. With tone 2, it means *spirit*.

In English, for example, accentuation changes the meaning in:

"record" which is a noun or a verb according to its accentuation on the first or second syllable.

The phenomenon of assimilation is also widespread. A phoneme sometimes assimilates a distinctive feature from a neighbouring phoneme.

Phonemes sometimes change one or more of their distinctive features according to their environment in the word and in the sentence (for example at the junction between words); for this reason, assimilation

