
A typology of language registers

Une typologie des registres de langues

STANDARDSISO.COM : Click to view the full PDF of ISO/TR 20694:2018



STANDARDSISO.COM : Click to view the full PDF of ISO/TR 20694:2018



COPYRIGHT PROTECTED DOCUMENT

© ISO 2018

All rights reserved. Unless otherwise specified, or required in the context of its implementation, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Fax: +41 22 749 09 47
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword	iv
Introduction	v
1 Scope	1
2 Normative references	1
3 Terms and definitions	1
4 General principles	3
4.1 Ability to vary and use language registers appropriately	3
4.2 Language registers and non-linguistic signals	3
4.3 Differences across languages	4
4.4 Descriptive and prescriptive registers	4
4.5 Text types, contexts, and linguistic characteristics	4
4.6 Language registers as formal to informal or high to low continua	4
4.7 Division into subregistries	5
4.8 Language registers and other language varieties	5
4.9 Languages used in document examples	6
4.10 Continuing evolution of language registers	6
5 Markers of different language registers	6
5.1 Lexical markers	6
5.1.1 General	6
5.1.2 Domain-specific vocabularies	7
5.1.3 Archaic words and phrases	7
5.1.4 Shortened and contracted form of words	7
5.1.5 Slang, swear words, profanities and taboo words	7
5.1.6 Use of foreign words and phrases	8
5.1.7 Register switch markers	8
5.2 Phonological markers	9
5.3 Grammatical features	9
5.3.1 General	9
5.3.2 Differences in morphology	9
5.3.3 Passive and impersonal constructions	9
5.3.4 Direct and active verb forms and use of specific tenses	10
5.3.5 Use of concise or periphrastic verb forms	10
5.3.6 Long/complex nominal phrases	10
5.3.7 Conjunctions and linking phrases	10
5.3.8 Sentence length	10
5.4 Honorifics and forms of address	11
5.4.1 General	11
5.4.2 Polite forms of address	11
5.4.3 Vocabulary and grammar in honorific forms of address	12
6 Mapping registers between languages	12
6.1 Register equivalence in different languages	12
6.2 Conveying language registers in translation	13
6.3 Inconsistencies in register	13
6.4 Loss of information in translation	13
6.5 Identifying language registers for translation memories	13
7 Language registers and diglossia	14
Bibliography	15
Alphabetic index of the terms and definitions	16

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 on 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 the following URL: www.iso.org/iso/foreword.html.

This document was prepared by Technical Committee ISO/TC 37, *Language and terminology*, Subcommittee SC 1, *Principles and methods*.

Introduction

This document aims to clarify ambiguities arising from the use of the term 'register' to designate different concepts. It aims to examine different conceptualizations of language registers amongst linguistic theorists so that useful definitions can be agreed on for use in, for example, standardization work. It also aims to contribute to developing new means of providing comprehensive coverage of language variants.

Computational management of language resources requires appropriate descriptors and tags for different language varieties.

A typology of language registers will aid appropriate communication in business and commerce, for example where a marketing campaign needs to address consumers in a friendly, informal register, or in medicine where there is a need to understand clearly the difference between technical communication between professionals on the one hand, and clear and simple communication for public health campaigns on the other.

There is as yet no common agreement on terms and concepts for individual language registers, or way to map their relationship to each other.

As dialects can have a geographic designation, language registers can be designated by their linguistic patterns, both lexico-grammatical and discoursal-semantic patterns associated with situations.

A description of a language register needs to state whether it is a written or a spoken register, or expressed by some other modality. It is therefore multifaceted, and polyhierarchical, fitting in with existing ISO standards such as ISO 24620-1 and ISO 639 in order to attain maximum impact. It lays the groundwork for developing codes for language registers at a future date. This document proposes an overview of all relevant language registers and does not intend to create standards for each individual register. It will, however, help to identify and inform further areas for elaboration.

STANDARDSISO.COM : Click to view the full PDF of ISO/TR 20694:2018

A typology of language registers

1 Scope

This document gives the general principles for language registers in both descriptive and prescriptive environments. It defines key concepts and describes examples of different language registers that can be applied across all or many languages and those that are language-specific. It lays down guidelines for the use of appropriate language registers needed in a wide range of environments. These include:

- terminology work, where it contributes to the development of a wide range of standards;
- translation, so that appropriate language levels can be chosen in target languages, to match that of the source language;
- lexicography, to improve descriptors of non-geographic language variants;
- second language teaching and learning, so that students can avoid pitfalls associated with inappropriate language use;
- software, to improve tagging of language variants in computer applications.

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.

3.1

language variety

largest subset of an individual language that is homogeneous both with regard to a certain criterion for linguistic variation and with regard to a certain structural criterion for linguistic variation

3.2

code switching

changing backwards and forwards between two or more languages or language varieties in conversation

3.3

language register

language variety (3.1) used for a particular purpose or in an event of language use, depending on the type of situation, especially its degree of formality

Note 1 to entry: An individual usually has more than one language register in their verbal repertoire and can vary their use of register according to their perception of what is appropriate for different purposes or domains.

3.4

common register

language register (3.3) used in addressing or referring to non-royal persons

3.5

royal register

language register (3.3) used in addressing or referring to royal persons

3.6

formal register

language register (3.3) that conforms to accepted standards and conventions and is used in serious and official situations

3.7

informal register

language register (3.3) that does not pay special attention to standards and conventions and is used in casual and familiar situations

3.8

high register

formal register (3.6) conceived of as being at the top of a vertical continuum of *language registers* (3.3)

3.9

low register

informal register (3.7) conceived of as being at the bottom of a vertical continuum of *language registers* (3.3)

3.10

slang register

extremely *informal register* (3.7) of a word, term, or text that is used in spoken and everyday language and less commonly in documents

EXAMPLE In aviation, the phrase *fly by the seat of your pants* is slang for the more formal *fly without instruments*.

3.11

facetious register

language register (3.3) related to an expression that is intended to be clever and funny but that is really silly and annoying

3.12

legal register

language register (3.3) used in the domain of law

3.13

literary register

language register (3.3) commonly used in literary works

3.14

neutral register

standard register

language register (3.3) appropriate to general texts or discourse

3.15

simplified language

prescriptive *language register* (3.3) following guidelines to make text and/or speech easier to understand

3.16

technical register

language register (3.3) appropriate to scientific texts or special languages

3.17

vulgar register

language register (3.3) of a term or text type that can be characterized as profane or socially unacceptable

3.18**in-house register**

language register (3.3) that is company-specific and not readily recognized outside this environment

Note 1 to entry: In-house terminology is not necessarily equivalent to bench-level terminology, inasmuch as the former can thrive at very high levels of research and development. In-house terminology is frequently the source of new technical terminology that eventually gains widespread acceptance on a broader scale.

3.19**bench-level register****shop term**

register of terms used in applications-oriented as opposed to theoretical or academic levels of language

3.20**subregister**

subsidiary division of *language register* (3.3)

3.21**honorific**

word, title, or grammatical form that signals respect or high social status

3.22**motherese****baby talk**

language register (3.3) often used by adults in addressing very small children

4 General principles**4.1 Ability to vary and use language registers appropriately**

Language registers are varieties of language where one person can have more than one type of language at their command, sometimes called their 'verbal repertoire'. Individuals are often able to vary their register according to different circumstances or environments which can depend on social attitudes, as well as geographic issues, which are covered by dialect and accent. Use of different language registers can be prescribed and proscribed for some environments, either through education and conformity to social norms, or, increasingly, through adherence to guidelines in various professional settings such as translation or publishing environments.

No individual has a perfect grasp of all possible language registers used in their language community. However, an individual's ability to understand a wide variety of registers and their social significance is greater than their ability to use, speak or write in these registers. Some higher registers are formal markers of learned environments, and lack of mastery of these registers is therefore taken as an indicator of a lack of education. Individuals from deprived backgrounds are sometimes characterized as having less mastery of different language registers, with this holding back their prospects of well-paid jobs. Thus, the active teaching of different language registers and the appropriate context for their use, is sometimes advocated.

The understanding and appropriate use of different language registers is also important in second language teaching, where errors of register mark students out as not being proficient even where their accent, vocabulary and grammar are exemplary. Understanding of register is also important in translation work, where the translator needs to translate into a similar register unless instructed otherwise. Different language registers also need to be kept separate in translation memories, otherwise texts of mixed language registers can result, leading to incongruity and lack of consistency.

4.2 Language registers and non-linguistic signals

The use of different language registers can be accompanied by the use of other non-linguistic signals, such as different gestures, dress, or even seating arrangements. Whilst these do not come under the remit of the current document, it is well to remember that language can rarely be considered in isolation.

4.3 Differences across languages

Some language registers are language-specific. In other cases similar language registers are found in many languages, especially where language communities share social characteristics, such as respect for certain members of society, or peer group solidarity. Language communities with very different social structures, for example a highly stratified community compared to one with no strong hierarchies, are more likely to find that some of their language registers have no equivalent in the other language. Different language registers can serve different purposes in different communities, for example to emphasise social distance, or to signal group membership.

4.4 Descriptive and prescriptive registers

A typology of language registers includes the descriptive, for example, of slang, informal varieties, literary registers, and the prescriptive, for example, simplified languages, and controlled languages for use in critical communications. From the perspective of a typology of language registers, simplified natural language is one specific example of a prescribed language register. Attempts to label different language registers are found in many descriptive dictionaries, although there can be inconsistencies in lexicographic descriptions. Guidelines on the use of prescriptive registers are sometimes published, for example, by the Plain English campaign, or as company or newspaper style guides.

4.5 Text types, contexts, and linguistic characteristics

This document follows the definitions in Clause 3 of language register and various subregisters. Different schools of linguists and literary theorists have understood terms such as 'language register', 'genre', 'text type' and 'style' in various ways, thus causing confusion and conflicting arguments. Although there is some overlap in these and other related concepts, in this document, 'language register' is related not only to a particular genre or text type, but is also identified by its linguistic characteristics. These internal linguistic characteristics make it possible to identify individual language registers, and develop computational methods of tagging them where needed, for example, in the management of translation memories where consistency of register is needed in the translation output.

Humans are usually able to vary the language registers they use according to contextual and situational parameters. In this language registers differ from dialects, which are associated with different groups of speakers with a geographical or societal distribution.

4.6 Language registers as formal to informal or high to low continua

There is some correlation between language registers and text type, with some sociolinguists arguing that language registers are linguistic varieties linked to occupations, professions or topics and are usually characterized solely by vocabulary differences. Others, however, point out that there is no general consensus within sociolinguistics on the use of register and related terms such as genre and style. Understanding language registers as only differences in vocabulary and closely tied to different domains is at odds with general language descriptions of 'high' and 'low' registers or 'formal' and 'informal'.

Examples of language registers include high/low registers, formal/informal registers, royal/common registers, neutral (standard) registers, technical registers, legal registers, mafia talk, slang, motherese (baby talk), clear and plain language. Language registers can be placed on a continuum from the most to the least formal, and from the most prescribed to the least prescribed use of language. Prescription can take the form of human-mediated guidelines, or computer-mediated control of language use.

The register continuum can be represented horizontally or vertically.

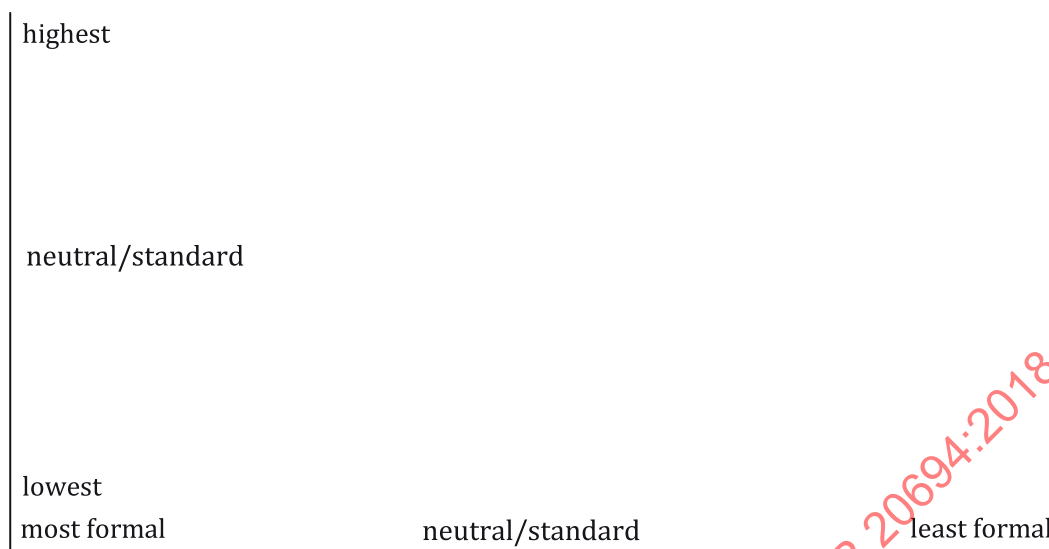


Figure 1 — Language registers on a vertical high to low continuum contrasted with a horizontal most formal to least formal continuum

Some registers, specifically royal ones, can represent a break in the continuum, where it is argued that they do not represent a 'yet more formal' register, but can be a form of diglossia, or a different kind of language with its own set of rules and conventions outside the normal range of language registers.

4.7 Division into subregistries

Individual language registers are further divided into subregisters, for example the technical register is often subdivided into domain-specific registers, such as medical register, legal register and engineering register. A further division of domain-specific registers can be in-house registers, which are confined to individual companies or organisations.

4.8 Language registers and other language varieties

The list of permissible values for the data category/register historically listed in the Data Category Repository (DatCatInfo¹⁾) as common in terminology databases does not conform to the categories of language registers discussed in this document.

Dialect is not considered a type of register in this document and is more appropriately classed as another type of language variety or category.

1) DatCatInfo is available at www.datcatinfo.net. It is maintained by LTAC/TerminOrgs. This information is given for the convenience of the users of this document and does not constitute an endorsement by ISO of this product.

Table 1 — Example of a possible typology

Register level	Names of registers in level	Other named registers
High	Royal register (Thai) Religious register Frozen or static register Formal-polite (Korean) Elegant speech (Japanese) Academic register	Formal register
Middle	Polite (Korean) Semi-formal or formal-lateral (Korean)	Neutral/standard register Bench-level register In-house register
Low	Informal register Casual register Familiar register Intimate register	Slang register Vulgar register
Unknown or variable level	Consultative register (could be middle or high) Trance register Email register	Dialect register Technical register Taboo register Ironic register Facetious register

4.9 Languages used in document examples

Examples in this document are primarily given in English. Translations of this document can provide examples in other language(s). This document does, however, also give examples in other languages when discussing features not usually found in English.

4.10 Continuing evolution of language registers

Language registers vary over time, with new language registers emerging due to new cultural and social contexts. Conversely, other language registers can fall out of use, for example, the old royal register previously used in Mandarin when China had emperors and a royal court is no longer employed. However, such language registers still need to be catered for and appropriately tagged when encountered in historical documents, films, works of fiction and similar contexts.

5 Markers of different language registers

5.1 Lexical markers

5.1.1 General

Examples of lexical markers are the vocabulary items used in different language registers. Searching for them and their frequency of use in text is an easy way to identify different language registers, especially using computational methods.

Lexical markers in English include use of the word 'child' in formal registers, contrasted with 'kid' in informal registers; 'periorbital haematoma' in a medical register, contrasted with 'black eye' in non-technical registers; 'urine' in formal and technical registers, contrasted with 'pee' or 'wee' in very informal registers or motherese. The use of different vocabulary items in different language registers is common across languages. Some words are marked according to usage in traditional dictionaries, with

register labels to alert readers that certain words are considered to belong only to formal, informal, archaic, literary, technical, humorous or slang registers.

5.1.2 Domain-specific vocabularies

Different subject fields each have their particular technical vocabularies, for example, medical domains will include technical terms which are specific to that domain, legal domains will also have domain-specific technical terms, as will engineering domains, etc. Inclusion, or comparative absence of, technical terms can therefore be used, in conjunction with subject-specific grammar, to identify and categorize text as belonging to the technical register.

Closely related subject fields can have different vocabularies if they have different communities of practice. Civil engineers differ from builders in their use of certain terms, for example using 'baluster' for the vertical units on timber stairs, whereas builders merchants can call them 'spindles', reflecting an older craft-based practice.

5.1.3 Archaic words and phrases

Certain words and phrases can be used in the most formal registers where they have dropped out of everyday usage. Examples in English include 'foresooth' to mean 'indeed', 'methinks' to mean 'I think', and 'thou' to address the second person singular, where contemporary English uses 'you' for both the singular and plural forms.

Fossilized expressions such as 'so be it', 'the more the merrier', 'long live the Queen' are usually used in higher registers, as is formulaic language. The use of prefabricated expressions is especially prevalent in the legal register, for example, 'cease and desist', 'draw up an agreement', 'unless otherwise stated'.

In some languages where older forms of the language have high prestige, such as ancient Greek for modern Greek speakers, archaic words and phrases can be very popular in more formal registers. This can lead to incorrect use, or to change via folk etymology to become part of the standard area. For example, in Greek, aigh-o-klima [αἰγιοκλίμα] 'goat-CM-vine' has become agi-o-klima [αγιοκλίμα] 'holly-CM-vine'.

In many languages also, archaic words and phrases can be used in a facetious register, for comic effect or to poke fun at people who try and use words and phrases that they think convey education and learning without fully understanding their meaning or appropriate context.

5.1.4 Shortened and contracted form of words

Shortened and contracted forms of words are markers of informal and familiar registers in many languages. Some of them, such as 'wouldn't' for 'would not' and 'don't' for 'do not' in English are accepted as standard in all but the most formal registers. On the other hand, contractions such as 'dunno' for 'I don't know' and 'c'mon' for 'come on' are considered as markers of very informal registers and are more prevalent in spoken modalities or social media. Online communication such as texting, email and instant messaging is especially prone to shortened forms and contractions, for example '2nite' instead of 'tonight', 'u' instead of 'you', and 'u2' instead of 'you too'.

5.1.5 Slang, swear words, profanities and taboo words

These are absent in all higher registers except where quoted, for example, in interview transcript for court cases. On the other hand they are strong markers of slang and vulgar registers, and appear also in other very familiar registers. The acceptability of use of swear words and profanities varies across languages and within regional language variants, for example, being more acceptable in Indian rather than in British English. Within some social media, the use of swear words and the vulgar register by private individuals is increasing as a way to differentiate between themselves and corporates using these media for publicity purposes.

5.1.6 Use of foreign words and phrases

5.1.6.1 General

Foreign words and phrases can be used in code switching or as regular borrowings.

5.1.6.2 Code switching

This is a phenomenon, common in bilingual and multilingual societies where words and phrases from one language are included in another language, more commonly in the spoken medium. In some language communities it is associated with informal registers, especially when young people make extensive use of international youth slang, both in written and oral texts. They can be a marker of in-group solidarity, mixed with slang and other low register phenomena.

Swear words are often used when code-switching, and many swear words can be borrowings from other languages, for example, many Greek swear words are borrowings from Turkish. The effect can be to lessen the impact of the swear words, thus raising the language register, for example in Polish where English swear words are more acceptable since they do not seem to have the strong communication effect that native swear words normally have. On the other hand, the impact of foreign swear words can be to lower the language register, for example in Welsh where English swear words are more offensive, because their meaning is well understood by the bilingual community, and where English borrowings are in any case markers of very informal registers.

In other language communities code-switching can also signal a switch to higher registers, where members of a community wish to signal their knowledge of a high-status language, for example as used by Thais. In South Africa, increasingly, code-switching, especially into English, is used when teaching to indicate a more academic register.

5.1.6.3 Borrowed words and phrases

These are high status words and phrases, often quoted as they are without being adapted to the host language. A certain set of foreign words and phrases can gain wide currency in the host language, at least within a certain social or professional group, for example, doctors, computer scientists. Some experts consider this different from code-switching, where there is greater freedom in the use of borrowed words or phrases. Others consider literary borrowings as a high status form of code switching. In English and other European languages they can come from Latin or French, for example, '*persona non grata*' and '*ad nauseum*' from the Latin, and '*sang-froid*', '*savoir-faire*' and '*bon voyage*' from the French. They are used in more formal registers and are sometimes considered as snobbish, especially if over-used in conversation.

Latin borrowings, for example, *a priori*, *de facto*, *sui generis*, are also used in technical contexts, indicating academic discourse and high registers in many languages. English is a contemporary source of borrowings in technical and scientific registers, due to the current hegemony of English in these domains. These borrowings, alongside native technical and scientific terms, can therefore be indicative of technical and scientific registers. However, in language communities where there is a prejudice against borrowing from a dominant language, and where there are efforts at terminology planning to counter domain loss, the use of borrowed technical terms and phrases can signal a more informal or a non-neutral register.

Borrowings also include words and phrases that have been adapted to the host language by means of changes to orthography, use of affixes, assignation of gender and other grammar features. Such borrowings are common in various registers and processes, including specialized communication and the creation of technical terms.

5.1.7 Register switch markers

Markers are sometimes used to signal a change of register. For example, in both Welsh and Irish, switching to English can be marked by an idiom meaning 'as the Englishman says'. In Greek, speakers

can mark a change from formal to less formal speech by drawing attention verbally to the switch, for example, by saying '*na to po laika*' [να το πω λαϊκά] 'to put it colloquially'.

Gestures such as air quotes (virtual quotation marks) can also be used, especially in English, to mark the introduction of lexical items more appropriate to another language register, usually to express irony, sarcasm or satire.

5.2 Phonological markers

In the spoken medium, words in more formal registers in many languages are enunciated more clearly and fully, with letters usually silent in less formal registers being pronounced, for example, the silent 'e' at the end of French words can be uttered in very formal speech. There can be other differences as well, such as in Greek, different consonant clusters used, for example *ptohos* [πτωχος] in a formal register vs *ftohos* [φτωχός] 'poor' in a neutral register; stress rules, for example *paráthirou* [παράθυρου] in a formal register vs *parathírou* [παραθύρου] 'window' in a neutral register; and vowel changes, such as i > e before -r-, for example, *ipiresia* [υπηρεσία] in a formal or neutral register vs *iperesia* [υπερεσία] 'service' in informal registers of some speakers, mostly elderly.

Change of stress placement can also be associated with particular registers. For example, in Polish words ending in *-ika* or *-yka* (such as names of sciences: *matematyka*, *fizyka*, *semantyka*) are pronounced with antepenultimate stress in careful speech and with penultimate stress in informal language. Similar variants are also observed in the case of certain past tense and conditional forms of verbs. It needs to be noted, though, that in Polish at least, individual speakers do not switch between the two possibilities available, but consistently stick to one, whatever register they are using, with better educated speakers using the more careful variants.

5.3 Grammatical features

5.3.1 General

Individual language registers are also identified by their use of different grammatical features. [Note: this document does not call these 'grammatical markers' as that term is already used in the different meaning of morphemes that indicate the grammatical functions of words and phrases]. Relevant grammatical features can differ widely between languages, but some are common to a number of languages. Often the most conservative grammatical features pertain to higher or most formal registers, while innovative and less standard features are indicative of lower or less formal registers.

5.3.2 Differences in morphology

Morphologically simplified forms are indicative of less formal registers in many languages. For example, declension and conjugation according to the *katharevousa* [καθαρεύουσα] morphological system is a high register feature of Modern Greek. Morphemes marked with the feature [+learned], which percolate to the final lexical structures, such as inflectional suffixes, for example, *taxe-os* [τάξεως] (high register) vs *taxi-s* [τάξις] (neutral register) 'classGEN'; derivational affixes or affixoids, for example, *epiplo-poieio* [επιπλοποιείω] (high register) vs *epipl-adiko* [επιπλάδικο] (neutral register) 'furniture-location/makeDS'; and stems as first or second constituents in compounding, for example, *pale-o-pol-eio* [παλαιοπωλείο] 'old-CM-town-locationDS' (high register) vs *pal-o-pedho* [παλιόπαιδο] 'old(>bad)-CM-town-child' (neutral register).

5.3.3 Passive and impersonal constructions

These are common in higher registers, especially academic discourse, and technical and scientific registers where they are used to provide an impression of objectivity. In English, the construction 'be' + past participle is often used, for example 'It was discovered that...' rather than 'We discovered that...'. However, writings in the medical register in English used to be typified by the use of the pronoun 'we', which is now spreading to other scientific subregisters.

Passive constructions are formed in higher registers of many languages by taking an impersonal pronoun as the syntactical subject. Examples of this are 'it' in English in the example above, 'il' in French, 'es' in German, 'het' in Dutch. In other languages, such as Welsh, there are specific impersonal forms of the verb, for example '*credir*' (it is believed), '*credwyd*' (it was believed). In Greek also, in cases where the agent is not specified, only the impersonal verb is used, for example, *pistevetai oti* [*πιστεύεται ότι*] (it is believed that).

5.3.4 Direct and active verb forms and use of specific tenses

These are indicative of many less formal registers, in both oral and written modalities. They are especially prevalent in less formal oral modalities where spontaneous speech is used, as opposed to prepared speeches and spoken texts. However, they are increasingly recommended for use where the subject matter is technical or scientific, in contrast to the traditional passive and impersonal constructions. Such recommendations are indicative of efforts to create an accessible language register, even for very technical and scientific subject matter, especially for English, where since it has become the international language of science, there is a need to cater for scientists and technologists whose first language is not English.

Plain English and simplified languages, where guidelines advise authors to identify the agent for the sake of clarity, are therefore examples of language registers that are lower on the formality scale but still high in terms of subject matter. Controlled language also belongs to this register, where a computer program is used to proscribe certain constructions, for example, with a message saying "do not use 'the oil should be changed', use 'change the oil' instead".

5.3.5 Use of concise or periphrastic verb forms

Some languages, for example French, Spanish, Welsh, have the choice of two or more competing methods of conjugating verbs, at least in some tenses. One method can be inflectional, while the other has a periphrastic construction, using an auxiliary verb to indicate tense. This is not common in English, but can be seen in English influenced by Welsh, for example 'I did go' instead of 'I went', 'I did do' instead of 'I did' where the verb 'do' is used as an auxiliary verb. In such languages, the inflectional form usually belongs to more formal, even archaic registers, whilst the periphrastic form belongs to less formal, or very informal registers.

5.3.6 Long/complex nominal phrases

A relatively high proportion of nouns in a text can indicate a scientific or technical register due to their use to convey ideas and to identify new objects or inventions. For example, noun compounds are multiword units formed by nouns modified by other nouns or adjectives. This is a common form of term formation in scientific and technical domains and is therefore a strong indicator of the attendant language register.

5.3.7 Conjunctions and linking phrases

More formal registers are associated with the use of longer and more complex conjunctions and linking phrases. In English, for example, 'as a result of which' can be paraphrased as 'and so' in less formal contexts. Other examples include the family of conjunctions and linking phrases associated with legal language, such as 'whereupon' or 'in witness whereof'.

5.3.8 Sentence length

Spontaneous speech is rarely delivered in long and complex sentences. Long sentences, and sentences with many subclauses, are, however, indicative of some literary, technical, legal and formal registers. Prepared speech in the form of public addresses, sermons, lectures, etc. can also exhibit similar features. Frequent use can be made of relative pronouns, for example, 'which, that, of which'; linking words, for example 'and, although, though, since, as'; prepositions, for example, 'despite, during'; and adverbs, for example, 'usually, meanwhile, firstly, secondly'.

However, such complexity is proscribed in plain language registers, and writers in some technical registers are now advised to keep their sentences short and simple in order to aid clear communication. Sentence length and complexity is therefore no longer a safe indicator of high registers.

5.4 Honorifics and forms of address

5.4.1 General

Many languages have different ways of addressing people according to politeness and their relative status, for example, subordinate to a superior, adult to child, employee to senior personnel, shop worker or waiter to customer, commoner to royalty. In some languages this extends beyond use of individual forms of address to use of different grammar, and even different vocabulary. These changes in an individual's use of language can be understood as different language registers, to be placed on the same formality/informality axis as other language registers, unless the royal register is so different as to be considered diglossic. Honorifics and differentiating forms of address are also usually accompanied by other markers of formality/informality in use, both in speech and in text.

Honorifics are more important and intricate in some languages compared to others, usually reflecting the culture of language communities. The culture of many Western language communities tend to place less emphasis on individuals' relative status, compared with many Asian and African language communities where the correct identification of status is an important part of both language and culture. In some languages such as Japanese, honorific/humble forms of address come as a contrasting pair where, for example in a customer/server relationship the server will use humble forms but will use the honorific when addressing the customer.

Failure to use appropriate honorifics and forms of address can cause serious offence in business, diplomatic and other circles. They need to be properly understood and observed also in third party reporting, for example, in news items, and in translating between languages. Less confident speakers can choose to communicate in a different language, such as English, where honorifics are not an issue, rather than cause offence by failing to use the appropriate honorifics.

It is sometimes thought to be comparatively easier to translate from a language with a rich and strict hierarchy of honorifics and many different forms of address into one with one with fewer honorifics and less differentiation in addressing people. In reality, honorifics and forms of address pose challenges for translators in either direction, as other, subtler means of indicating politeness, social stratification and closeness/distance must be sought where honorifics and explicit linguistic markers are insufficient.

5.4.2 Polite forms of address

In many languages, the second person plural form of address is also used as a polite form to show respect instead of the second person singular. This was true of older forms of English, where 'thou' was the second person singular form, and 'you' was the second person plural form. However, increased use of the second person plural form as a form of address in the singular means that the use of 'thou' is now restricted to archaic use and certain dialects. In some geographic areas and registers, new plural forms of address is emerging, for example, 'you(s) guys, y'all, you lot'. However, these forms are not yet accepted in Standard English, and while they can be used as indicators of informal registers, do not indicate politeness or lack thereof.

Other indicators of politeness include the way names and titles are used, both in addressing people directly, and in talking about them in the third person. Although certain conventions in the use of names and titles vary across languages, in general, first names are used in informal registers, for example, 'Hi, John', compared to using a title plus surname in formal registers, for example, 'How are you Mr Jones?' The use of full names, without titles, can also indicate formal registers, for example, 'I was given this book by John Jones'. In some technical registers, such as academic writing, the surname on its own is used, for example, 'Jones (2016) said...' In informal registers, terms of endearment are often used, for example, 'How are you, love?' In very formal situations, a title can be used without a personal name, for example, 'Sir' or 'Madam'. Formality can be indicated in English by the use of titles such as 'Mr

President' or by addressing someone by the name of their profession, for example, 'Doctor, Professor, Minister, Waiter'.

In such examples it is expected that the choice of forms of address is in keeping with other indicators of language register used in the exchange, i.e. a formal form of address would be accompanied with other indicators of formality in a consistent manner throughout. The exception is when formal forms of address are used in a jocular manner as when teasing, or in a facetious register, for example 'Well madam, don't you look half grand today'.

5.4.3 Vocabulary and grammar in honorific forms of address

In languages which place great emphasis on the correct identification of status, and the corresponding use of appropriate honorifics, different vocabulary and different grammar can be involved in the different language registers, as well as different forms of addressing different people. For example, Japanese has three broad levels of politeness: casual, polite, and honorific/humble which can be classed as different language registers. Some vocabulary items differ according to register, especially in the honorific/humble register where different words are used for some concepts, for example, the words for 'room, bathroom, customer'. Verb conjugations can also have different forms for the honorific and both nouns and verbs can use honorific prefixes to denote register.

Similar vocabulary and grammar differences according to honorific register apply to other Asian languages such as Korean and Tibetan. Pronouns, nouns, verbs, adjectives, adverbs and auxiliaries can variously be affected in such languages, according to conventions of status and politeness. Although register and dialect are usually considered to be two different dimensions, the use of the honorific register can be more common in some Tibetan dialects than others. For example, in Tibetan, it is more common in the standard language and the dialects of Central Tibet.

6 Mapping registers between languages

6.1 Register equivalence in different languages

Many language registers are specific to individual languages, while others exist across more than one language. A number of languages share the conceptualization of language registers along a high to low axis, or from most formal to the least formal, and some also use similar appellations, for example 'classical', 'technical', or 'slang' to denote individual registers. Common patterns in the markers of individual registers can also be found across languages, including:

- more conservative linguistic features at the most formal side of the formality axis, contrasted with more innovative linguistic features at the least formal side of the formality axis;
- greater adherence to grammatical and lexical norms at the most formal side of the formality axis, contrasted with increasing disregard for grammatical and lexical norms at the least formal side;
- in speech, clearer and fuller diction at the most formal side of the formality axis, moving towards contracted forms and less clear diction at the least formal side;
- more reference to international standards and conventions in technical registers, together with conscious efforts to align to a mappable common scientific discourse, contrasted to inward (language-specific) references in other registers, often with appeal to linguistic purism in more formal registers, and conscious or unconscious influences from other contact languages in less formal registers.

Thus it is possible to speak about language registers such as 'formal register', 'technical register' or 'slang register' in many languages, and they will share some common features, whilst also exhibiting many linguistic markers of language register that belong only to individual languages.

6.2 Conveying language registers in translation

Translation is primarily focused on conveying meaning from one language to another. However, translators will usually strive to carry over other aspects of the source language, including the specific language register(s) used in the original, into the target language, unless specifically instructed to do otherwise, for the sake of the intended audience.

An example of instructions not to follow the language register(s) of the source languages is when a translator is asked to simplify the language and use a non-technical register, where the source text was written in a higher or more technical register.

Creative texts such as novels or film scripts can contain many different language registers in order to portray different characters, situations and relationships. Replicating these complexities of language register in the target language adds to the success of the translation and the enjoyment of the translated product. Less complex texts, such as company documents and reports, tend to adhere to one language register only.

The choice of an equivalent, appropriate language register or registers for the translation in the target language is an important one. For example, translating a carefully crafted, formal speech in the source language into a lower colloquial register in the target language could interfere with the gravity of a serious message, or translating a publicity leaflet written in a simple, clear register in the source language into a high, formal register in the target language could make the message inaccessible to the intended audience.

6.3 Inconsistencies in register

Unless change of language register within a text is a deliberate choice to mirror register change in the source, inconsistent use of language register is a feature of poor quality translation, and should be avoided.

Inconsistent use of language register is also a feature in the speech and writing of learners of another language. This can include inappropriate use of honorifics, mixing formal and informal lexical items and grammatical features, even to the extent of trying to combine elements within words, for example, formal stems with informal affixes.

Teaching the correct use of different language registers is therefore important, both within translator training and the teaching of foreign languages, and should form a core part of the student's curriculum.

6.4 Loss of information in translation

Translating from languages with a rich hierarchy of registers into languages with less choice of registers (situation of convergence) can be challenging. Subtleties of meaning and attitudes can be lost, for example, translating from a language such as French which has two forms of address in the singular, one informal (*tu*) and the other more polite and distant (*vous*) to a language such as English which has only one form to cover all (*you*). In this case, information about the relationship between the addresser and the addressee is lost, and must be surmised from other clues of context. Choosing the appropriate register to convey the correct degree of formality can include using other markers of register, such as choice of vocabulary and grammatical forms that are consistent with the desired outcome.

Conversely, translating from a language with less choice of register into one with a richer hierarchy of registers (situation of divergence) can present problems, for example in choosing the correct register for addressing a company director in an East Asian language with very narrowly defined rankings of politeness, and how would this carry over, not only to direct speech but to all other vocabulary, grammar and syntax issues in the translation project.

6.5 Identifying language registers for translation memories

Better results when using translation memory are achieved when texts stored and retrieved from the memory are similar in nature. This has led some users of translation memory software to classify their