

Correspondence to:

Translation Section Editor
Gabriele Berghammer,
the text clinic,
Vienna, Austria
gabi@the-text-clinic.com
www.the-text-clinic.com

Gained in translation: science at the multilingual crossroads

The spectre of translation quality – Part I

Quality matters in medical translation

In recent years, 'translation quality' has become a buzzword in the translation industry. Particularly since the introduction of European standard EN 15038¹ in 2006 and the certification process that has come with it, many translation service providers (TSPs) have been advertising their proprietary quality management methodologies as a guarantee for success.

But what is EN 15038, and – perhaps equally important from the point of view of quality assurance – what is it not?

In brief, EN 15038 regulates the requirements for translation services and creates a general framework for the interaction between clients and service providers in terms of each party's rights and obligations. Thus, the standard is exclusively concerned with setting up a standardized translation process and implementing measures designed to create a sustainable working environment. Importantly, however, EN 15038 is silent as to how to actually assess the quality of the end product arising from the translation process, i.e. the translated text.

While having a sound process in place is certainly an important prerequisite for delivering high-quality output, it is not in itself already a measure, let alone guarantee, for product quality. Alongside EN 15038, therefore, some TSPs have developed proprietary quality-assessment metrics designed to measure the quality of translated text. Some of these metrics are reportedly based on SAE J24502, the only standard so far available for rating the quality of translation deliverables. Overall, however, such metrics, while spotting the more obvious shortcomings in a text, such as wrong meaning or terminology, omissions, additions, or punctuation errors, fall short in assessing a

translated text for style or register, making them 'unsuitable for evaluations of material in which style is important'.²

But more on these standards later. Let us first consider some of the reasons why quality in medical translation – as indeed the quality of any text written in a field as sensitive as medicine – is important at all.

Quality matters because...

I see three main reasons for why quality in medical translation matters. First, the requirements for medical texts are that they be error-free. If they are not, they have the potential to cause serious harm or even death. Second, scientific texts should be easily readable and unequivocal. If they are not, they may confuse or mislead. Finally, translated scientific texts should mimic the style characteristics of the text genre in question in order not to make the text sound awkward, thereby undermining the credibility of the author of the source text.

... translation errors in medicine can be dangerous

Accurate and readable instructions for drugs or medical devices may be as important a safety issue as adequate hygiene in the operating theatre.³ While statistics about how often translation errors actually do cause harm are not available, some reports suggest that the danger is real.

In 2004, Mead Johnson Nutritionals had to recall two different baby food products because the instructions on how the products were to be prepared had been incorrectly translated from English to Spanish. Both the 16-ounce powder infant formula and the 32-ounce ready-to-use infant formula had dangerous preparation instructions, according to the US Food and Drug Administration (FDA). It reported that, if the baby food were prepared according to the incorrect Spanish instructions, the formula could cause seizures, irregular heartbeat, renal failure, and death.⁴

The importance of translated product labelling was also highlighted by a much publicized case from

Berlin, where 47 patients having had knee replacement in 2006 and 2007 had to undergo re-operation because physicians had implanted the knee prostheses without applying the necessary bone cement.⁵ The manufacturer had shipped the device without German instructions for use. Because the English phrase ‘non-modular cemented’ on the package of non-modifiable prostheses requiring cementing had been taken to mean ‘not requiring cementing’, hospital staff had sorted the cemented prostheses into the shelf for cement-free prostheses, and patients received prostheses that should have been cemented but were not. The error was not noticed until the US manufacturer started shipping the product with German-language stickers on the outer carton.

A 2007 literature review performed to identify papers on translating clinical and medical research documents identified only 44 relevant articles.⁶ Ten of the 44 articles described error types arising during translation, with an inability to obtain cultural equivalence and oversimplification of crucial information the most frequently mentioned sources of error. Unfortunately, the documents reviewed said nothing about the frequency of errors in medical translation, and many, in fact, dealt with interpreting rather than translating.⁷

It is likely that only a fraction of translation errors ever become public. For example, I once coordinated the translation of the Summary of Product Characteristics (SPC) for a medicinal product authorized via the centralized procedure. Requiring translation into multiple languages, the project was outsourced to a major TSP specialized in medical and medical device translation. The product in question was a solution designed for subcutaneous injection. The German translation returned by the TSP, instead of translating ‘administering’ or ‘injecting’ the solution as *verabreichen*, *anwenden*, or *injizieren*, used *einnehmen* throughout the entire text, suggesting that the drug be ‘swallowed’ or ‘taken orally’. This (and other, similar, errors) were spotted early enough in the review process to not actually cause confusion or harm – but I was surprised that such an error could occur at all, considering that the TSP reportedly not only employed expert translators but also had rigorous quality assurance (QA) procedures in place.

Alternatively, errors that do not get caught in time may go unnoticed because they are mentally amended by the reader who, even though faced with a text that contains an error or is equivocal, corrects it to mean what he or she knows (or thinks) it should mean.

From the few reports that do get publicized, it is difficult to determine where a translation error actually originated.

There are what may be referred to as ‘intrinsic factors’⁶ influencing the quality of a translation, referring mainly to the qualification and subject-matter knowledge of the translator. Thus, errors may arise from a lack of proficiency and medical background knowledge of the translator. They may also be due to instances of oversight by the experienced expert translator – an error category which, just as human failure in other areas of life, will be difficult to eliminate altogether. In medicine, inadvertently misplacing a comma can have disastrous consequences.

Then there are a number of ‘extrinsic factors’ influencing the quality of medical translation. As the examples above illustrate, these include a lack of awareness on the part of the manufacturer or marketer of the importance of making documents in a client’s native language available, with either no translation provided at all, the translator not given enough time or resources to do a proper job, or some other process-related shortcoming that precludes even a proficient translator from delivering a high-quality product.

Overall, a combination of well-versed translators and vigorous QA procedures, including an effective review process, may be expected to reduce the number of ‘critical’ translation errors, i.e. errors potentially leading to patient harm, to a minimum. However, there may be other sources of confusion or misunderstanding resulting from poorly written, imprecise, or misleading phrasing.

... readability is a sine qua non in medical communication

No matter how technical or non-technical a document may be, it does not serve its audience unless it is easily understandable, i.e. readable. Writing is not readily comprehensible when it is impossible or difficult to interpret, takes too long to make the point, or uses imprecise language. For sentences to be readable, they should use correct grammar, punctuation, and spelling. However, correctness alone is not a guarantee for readability.

In general terms, our writing style depends on the words we choose, the length of our sentences, the way we connect them, and our tone and register. A readable text is consistent, i.e. it uses the same key terms for key concepts and the same spelling and other linguistic and typographic conventions throughout. A readable piece of writing is clear, i.e. any one sentence requires no more than a single reading. A readable text uses exact wording, i.e. words and phrases that communicate rather than obfuscate. A readable text is concise, i.e. it conveys the most information in the fewest words without omitting details. It is fluent, i.e. easy to

read because of clear connections, variety, and emphasis. And it is 'graceful'.⁸⁻¹⁰

The readability of health-related texts has been given some scrutiny in the scientific literature. For example, a US study published in *Pediatrics* in 2003 found that installation instructions for child safety seats generally exceed the reading skills of most consumers, leading to improper installation.¹¹ Motor vehicle collisions, the authors explain, are a leading cause of death in infants and children, and the single strongest risk factor for injury in car accidents is the non-use of a restraint.

A study on the readability of patient information regarding breast cancer prevention from the website of the US National Cancer Institute also found that the information was written at far too high of a reading level.¹² Also, it has been shown that patients considering to participate in a clinical study may often not be able to give valid consent because they do not understand the study as a result of the low level of readability of the information material they are given.¹³

According to current legislation, the information in package leaflets for medicinal products must be easy for patients to read and understand. A Spanish study analysed the readability of the package leaflets of medicinal products through application of the Flesch formula, selecting the 30 medicinal products most widely consumed and the 30 which generated the highest expenditure during 2005 in Spain. Only five documents obtained an acceptable Flesch score, i.e. a score of 10, while 18 scored 0 and half of the documents had values below 2.¹⁴ Poor readability has been shown to lead to patients becoming fatigued and discouraged, which may affect compliance.

Inefficient or inadequate style makes readers work harder than they should. Writing a clear text is the author's responsibility. The reader's job is merely to follow the author's thinking and – depending on the text type – agree or disagree; the reader's job is not to 'decode the text'.¹⁵

If these aspects are common requirements for readable text, one would expect the same principles to apply to translated text. There is a close relationship between translation and writing. Translation 'may be looked upon as framed writing, obeying the same rules within a specific framework defined by the original'.¹⁶ As Didaoui has noted,¹⁶ 'rules governing translation as a text-producing exercise are basically the same as original textualisation', taking into consideration any shifts required as one language is transposed into another. Didaoui even goes so far as to state that "the word 'translation' may even be substituted with 'text-producing in the target language'".¹⁶

The consequences of bad writing can be grave: at its worst, the writing can become unethical, namely when it confuses or misleads. At the very least, it can become less powerful or persuasive. Or, in the case of medicinal products, it may delay marketing authorization. In 2009, according to the record of a telephone conversation between FDA and GlaxoSmithKline (GSK) Biologicals regarding a new submission of one of the company's vaccines, FDA's Center for Biologics Evaluation and Research (CBER) requested a number of Standard Operating Procedures (SOPs) to be submitted by the applicant and that these be made available 'within 2-3 weeks'. GSK stated that the SOPs would have to be translated into English, but that the translations should be available to CBER 'within several days'. SOPs are highly complex documents that portray a company's entire research & development and QA process and that usually take a long time to compile and finalize, and the initiated translator will wonder how any such documents could be translated within a matter of 'several days'. Indeed, in the telephone report, CBER states that there 'appear to be a number of translation errors. The SOP instructions are not clearly written'.¹⁷

Revising or correcting translated texts that lack clarity and readability can range from simple to tedious. Some sections can be improved by simple editing, as the following sentence with an ambiguous referent shows:

Treatment of infections with dermatophytes with terbinafine is a good option in transplant recipients.

Turning 'dermatophyte' into an adjective improves the readability rather effortlessly:

Treatment of dermatophyte infections with terbinafine is a good option in transplant recipients.

In the next example, the rather long list of nominal groups may require a second reading:

Patients were eligible for inclusion into the study if they required treatment with prothrombin for acute bleeding, overdose of coumarin or coumarin derivatives or prophylaxis.

With a comma of separation added before the last noun, one reading will suffice:

Patients were eligible for inclusion into the study if they required treatment with

prothrombin for acute bleeding, overdose of coumarin or coumarin derivatives, or prophylaxis.

The next sentence is derived from the German translation of the SPC of a centrally approved vaccine.

Die erste Dosis kann ab Vollendung der 6. Lebenswoche, sollte jedoch nicht später als vor Vollendung der 12. Lebenswoche verabreicht werden.*

Whereas the first part of the German sentence is unequivocal, the phrase *nicht später als vor Vollendung der 12. Lebenswoche*, a sort of literal translation from English, leaves the reader puzzled. Both readability and accuracy are enhanced by simply deleting the words *nicht später als*:

Die erste Dosis kann ab Vollendung der 6. Lebenswoche, sollte jedoch vor Vollendung der 12. Lebenswoche verabreicht werden.

Sometimes, however, simple editing may not be enough to improve readability, and not all QA procedures appear to consider readability an important textual feature. The next sentence is again derived from the German translation of the vaccine SPC introduced above:

Bei 5.673 geimpften Säuglingen (2.834 Säuglinge erhielten den Impfstoff) wurde die Wirksamkeit anhand der Abnahme der Inzidenz RV-bedingter Gastroenteritis durch die Impfstoff-G-Serotypen (G1 bis G4), die frühestens 14 Tage nach Gabe der dritten Dosis [des Impfstoffs] auftraten, über die gesamte erste Rotavirus-Saison nach Impfung gemessen.**

With complex sentences such as these, simple editing is unlikely to enhance readability. The phrase would have to be recast in German, departing from the syntax of the English source and disentangling the nested sentence structure, such as in:

Bei 5673 geimpften Säuglingen, von denen 2834 der Impfstoffgruppe zugewiesen worden

*English original: 'The first dose may be administered from the age of six weeks and not later than the age of 12 weeks.'

**English original: 'In 5,673 vaccinated infants (2,834 in the vaccine group) protective efficacy was measured as a reduction in the incidence of rotavirus (RV) gastroenteritis caused by vaccine G serotypes (G1-G4) that occurred at least 14 days after the third dose of vaccine through the first full rotavirus season after vaccination.'

waren, wurde die Schutzwirkung anhand der Abnahme der Inzidenz Rotavirus (RV)-bedingter Gastroenteritiden durch die Impfstoff-Serotypen G1-G4 erhoben. Der Beobachtungszeitraum erstreckte sich dabei von Tag 15 nach Gabe der dritten Dosis [des Impfstoffs] über die gesamte erste Rotavirus-Saison nach Impfung.

...inadequate translation undermines credibility

If a translated text is neither wrong nor misleading, it may still sound awkward. This may be less of a practical danger, but it is potentially harmful to the author's reputation. When the writing, or the translation, is sloppy – what reason does the reader have to believe that the quality of the research the text describes is not?

[...] the medical profession (particularly clinical medicine) is full of jargon and idiosyncratic phrases which sound unusual, to say the least, in the context of everyday speech or writing. [...] The temptation may be great to change or omit these often awkward-sounding phrases, but they are so much a part of the professional language that the translator who does so is actually making a radical change in the register of the text; and to medical ears, the text becomes jarring and sounds 'less professional' without these familiar phrases. Not only does this make it more difficult for the medical professional end-user to quickly grasp the substance of the communication, but I believe it also has the undesirable effect of undermining the scientific credibility of the article or text (even if only subliminally).¹⁸

Biomedical communication does have a distinct style – or, rather, distinct styles – and these should be mimicked in translation, requiring an immersion into a particular discipline to appropriate its language.

Specialized language serves a specific purpose that cannot be accomplished either by the use of general language or by the specialized language of another discipline.^{19,20} Therefore, with writing being bound by the conventions of a particular genre 'one writing doesn't fit all'.²¹ These insights are far from new. The Roman rhetorician Quintilian said that every piece of writing requires 'a different and distinct style. [...] Every species of writing has its own prescribed law, each its own appropriate dress'.²²

Scientific language is intricately linked with the way scientific knowledge is generated, and this may be different in different areas of scientific

research and at different points in time. In this context, Thielmann²³ mentions two aspects that Ehlich has shown as characterizing scientific communication, namely that scientific texts are designed for a communicative situation in which any new finding is *a priori* considered controversial and has yet to be ratified by scientific peers. Also, the linguistic inventory of the language of science cannot be grasped on the basis of a purely terminology-oriented analysis, with many of the phrases used in science communication reflecting individual aspects of the cognitive process prevailing in science (e.g., *einen Grundsatz ableiten, eine Erkenntnis setzt sich durch*).²³

Therefore, translators will have to analyse the language and style of the source text and find an equivalent in their target language. The challenge of translating is not only to transpose scientific content but also to adapt the source-language 'dress code' to conform with the conventions expected by the target-language reader.

Translation: industry or craft?

Translation requirements are increasing worldwide, probably as a direct consequence of globalization. Neither the drive towards globalization nor the need for translation is new. For centuries, societies have striven to expand their spheres of influence through colonization or conquest, marrying and giving in marriage*, buying and selling. And throughout history, translation has been a loyal companion facilitating international communication.

What is different today is the speed globalization has gathered in the past two decades, largely as a result of technological advances that have compressed, or 'annihilated', space and distance.²⁴ Today, global companies bring their products to multiple markets at virtually the same time. Translation is a vital prerequisite for industrial internationalization and, aided by numerous software tools and applications and involving diverse experts from project and terminology managers to computer programmers, editors, graphics designers, and desk-top publishing specialists, has itself become an industry. At the same time, however, translation proper – the process of transferring a text across culture barriers – continues to be an intellectual activity that defies industrialization and requires know-how, expertise, and a human brain capable of anatomizing a source-

language text and sewing it back together in the target language. Keeping this in mind, it becomes clear that even the most robust translation process, unless relying on expert translators who master their craft, will not necessarily bring forth a target text that is error-free, readable, and, may I say, graceful – concepts which, admittedly, have yet to be defined.

References

1. Deutsches Institut für Normung [DIN]. Translation services – Service requirements; English version of DIN EN 15038:2006–08. Berlin, Germany: Beuth Verlag GmbH; 2006.
2. SAE J2450 Translation Quality Metric Task Force. Translation quality metric – surface vehicle standard. Available from: http://standards.sae.org/j2450_200508, 2005.
3. Crimson Life Sciences. Medical device labeling survey by Crimson Life Sciences exposes serious translation risks, Reuters, 19 March 2009.
4. European Business Forum. Mead Johnson Nutritionals product recalls (Brief Article) HighBeam Research. (2010 March 4). Available from: www.highbeam.com/doc/1G1-134816639.html, 2004.
5. Honert M. Klinik setzte 47 Patienten falsches Gelenk ein. Der Tagesspiegel, Available from: www.tagesspiegel.de, 11 August 2007.
6. Garcia-Castillo D, Fetter MD. Quality in medical translations: a review. J Health Care Poor Underserved 2007;18:74–84.
7. National Standard Guide for Community Interpreting Services. Available from: www.multi-languages.com/materials/National_Standard_Guide_for_Community_Interpreting_Services.pdf. Toronto, Canada: Healthcare Interpretation Network (HIN); 2011.
8. DuBay WH. The principles of readability. Available from: www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/bf/46.pdf, 2004.
9. Huth EJ. How to write and publish papers in the medical sciences. Baltimore, London, Munich: Williams & Wilkins; 1990.
10. Lannon JM. Technical communication. New York, Madrid, Amsterdam: Longman; 2000.
11. Wegner MV, Girasek DC. How readable are child safety seat installation instructions? Pediatrics 2003; 111:588–91.
12. Hoppe IC. Readability of patient information regarding breast cancer prevention from the web site of the national cancer institute. J Cancer Educ 2010;25:490–2.
13. Knapp P, Raynor DK, Silcock J, Parkinson B. Performance-based readability testing of participant information for a Phase 3 IVF trial. Trials 2009;10:79.
14. Barrio-Cantalejo IM, Simon-Lorda P, March Cerda JC, Prieto Rodriguez MA. [Grammatical readability of the package leaflets of the medicinal products most widely consumed and generating the highest expense in Spain during 2005]. Rev Esp Salud Publica 2008;82:559–66.
15. Zeiger M. Essentials of writing biomedical research papers. McGraw-Hill Higher Education; 1999.
16. Didaoui M. Translation quality: a pragmatic and multidirectional approach. In: Schmitt PAJH (ed.). Translationsqualität. Leipziger Studien zur

*Referring to the motto of the Habsburgs to have their members marry into other royal families to forge alliances: *Bella gerant alii, tu, felix Austria, nube*. 'Let others wage wars, but you, happy Austria, marry'.

- angewandten Linguistik und Translatologie ed. Frankfurt: Peter Lang; 2006.
17. FDA. Record of Telephone Conversation, Hiberix. Available from: www.fda.gov/BiologicsBloodVaccines/Vaccines/ApprovedProducts/ucm182111.htm, 2009.
 18. O'Neill M. Who makes a better medical translator: the medically knowledgeable linguist or the linguistically knowledgeable medical professional? A physician's perspective. *Am Translators Assoc Scholarly Monogr Ser* 1998;X:69–80.
 19. Waldo ML. Demythologizing language difference in the academy: establishing discipline-based writing programs. Mahwah, NJ: Lawrence Erlbaum Associates; 2004.
 20. Bourelle A. Lessons from Quintilian: writing and rhetoric across the curriculum for the modern university. Available from: www.worcester.edu/Currents/Archives/Volume_1_Number_2/CurrentsV1N2BourelleP28.pdf. *Curr Teaching Learn* 2009;1.
 21. Hall M. Book review on “Demythologizing language difference in the academy: establishing discipline-based writing programs”. *Composition Stud* 2005; 33:137–40.
 22. Quintilian. *Institutio oratoria* (translated by Murphy JJ). Carbondale, IL: Southern Illinois University Press; 1987.
 23. Thielmann W. *Deutsche und englische Wissenschaftssprache im Vergleich*. 3rd ed. Heidelberg: Synchron Wissenschaftsverlage der Autoren; 2009.
 24. Stanford Encyclopedia of Philosophy. Available from: <http://plato.stanford.edu/entries/globalization/>, 2011.

Gabriele Berghammer
the text clinic
www.the-text-clinic.com

Google translation

It can be quite handy to pop a simple French or German text into Google translation if you just want to get the general gist of what it means in English. Google is not great at tenses, e.g. I found that text written in the past tense in German was translated into the past perfect in English. And Google cannot be accused of lacking fantasy. Here is an example.

Original German: Ende Juni, Anfang Juli bin ich dann eine Woche nach Kroatien zu meinem Bruder gefahren. Er hat dort einen Wohnwagen in einem Nudistencamp stehen.

Google translation: Late June, early July, I'm a week after Croatia to my brother is run. The bear has a caravan in a nudist camp standing.

My translation: End of June, beginning of July I then went to Croatia to my brother for a week. He has a caravan standing in a nudist camp there.

I have translated *bin gefahren* as *went* but the literal translation would be *drove*. There is no hint



© Anders Holmqvist, 2012

of 'run' nor any sign of bears in the original German text.

Elise Langdon-Neuner
editor@emwa.org

Concepts from the linguistic crossroads

What's in a word...?

Ever thought about what a word is? In rather technical terms, a word may be defined as 'a sequence of letters with an orthographic space on either side'.¹ Taking a more philosophical stance, a word is 'the smallest unit of language that can be used by itself'² and that has literal (semantic) or practical (pragmatic) meaning.

We tend to think of a word as the very element in a language that carries meaning. Yet, meaning can be carried by units smaller than a word – morphemes. A morpheme cannot be further broken down into other elements of meaning and very often cannot be used on its own. For example, the morpheme 're', such as in 'rebuild' or 'recapitulate', means 'again', and cannot stand alone. The morpheme 'hyper' in 'hypersensitivity' means 'excessive' and is also used in composite words although it has, since the early 1940s,³ also been used as a stand-alone colloquial shortening of 'hyperactive'.

cross-road *noun* 'kro\s-rōd also -'rōd\
 a: the place of intersection of two or more roads
 b: (1) a small community located at such a crossroads (2) a central meeting place
 c: a crucial point especially where a decision must be made⁴

Morphemes can have a grammatical function, e.g. the suffix 'ity' in 'hypersensitivity', where it forms an abstract noun from an adjective. Also, morphemes may be used to form a plural (texts) or a tense (reported) or to turn an adjective into an adverb (hyperactively).

Why would this be of relevance for translation? Because very often there is no one-to-one relationship between word and meaning in different languages. In isolating languages, such as Vietnamese, there is a one-to-one correspondence of morphemes to words, i.e. any one word contains only one morpheme. By contrast, the two-morpheme English word 'disbelieve' is represented by two German words, i.e. *nicht glauben*,

and the German *Handrücken* is 'dorsum of the hand' in English. Overall, therefore, an element of meaning represented by a single word in one language may be represented by a number of words in another.

In the language of medicine, many terms are made up of Greek or Latin roots, but they may also originate from common speech. The same register in different languages may make different use of these Greek, Latin, and common-speech roots. For example, the English 'metacarpals', made up entirely of Greek morphemes, is *Mittelhandknochen* in German, consisting of common-language morphemes only. The Greek-derived term *ephelides* finds its English equivalent in the Scandinavian-derived two-morpheme word 'freckles', which in German becomes the three-morpheme *Sommersprossen*, a word which also highlights an additional aspect of meaning, namely that freckles, or 'summer sprouts', appear on the skin when exposed to the summer sun.

In translation, words may pose a problem when they refer to culture-specific concepts, such as the English 'copyright' and the German *Urheberrecht*, which, although very often used interchangeably, have rather different meanings. The English morpheme 'copy' derives from the Latin *copiare*, meaning 'to write in plenty' or 'to write an original text many times'³ and placing the emphasis on who holds the right to reproduce or commercialize a piece of intellectual property. The German morpheme *Urheber* derives from the Old High German *urhab*⁵ and focuses on who 'brought into being' or 'created' a piece of intellectual property. The difference in meaning between the two composites, therefore, should not come as a surprise.

More often than not, of course, meaning is carried by structures larger than a single word.

References

1. Baker M. In other words. A coursebook on translation. London, New York: Routledge; 2010.
2. Bolinger D, Sears D. Aspects of language. New York: Harcourt Brace Jovanovich; 1968.
3. Harper D. Online Etymology Dictionary. Available from: www.etymonline.com/
4. Webster's International Dictionary. Available from: www.merriam-webster.com/
5. Digitales Wörterbuch der Deutschen Sprache. Available from: www.dwds.de

LINGUEE has come of age

www.linguee.com

The web service Linguee – the search engine combing the internet for translated texts and making them available as a bilingual data pool that can be searched for words and phrases – has truly come of age. After a 1-year beta testing phase, the full version of Linguee, the then German-English bilingual translation tool, went live in May 2010. Since then, Linguee has expanded its service to include English-Spanish, English-French, and English-Portuguese as additional language pairs and has come to rank among the top 100 websites in Germany.

A specialized computer program – a web crawler – automatically searches the internet for webpages containing bi- or multilingual content. The texts are evaluated by a machine-learning algorithm, and translated sentences and words are extracted. The system is capable of autonomously learning to filter out the best translations based on quality criteria continuously refined on the basis of user feedback. Of the more than a trillion sentences that Linguee computers have already compared, only the top 0.01%, i.e. 100 million of the translated sentences, have been retained.

Linguee presents words in context

One major advantage over traditional dictionaries is that Linguee presents any word or phrase in the context of an entire sentence.

Many of the texts Linguee is based on derive from European institutions or EUR-Lex, the database of

EU legislative texts. For example, some of the text pairs displayed when looking for German ways of translating the phrase ‘application for marketing authorisation’ are displayed in Figure 1.

Linguee provides direct access to the source texts

A really nice feature of Linguee is that it does not only display translated sentence pairs, but also takes you straight to the documents the translation derives from. For example, clicking the ‘eur-lex.europe.eu’ hyperlink in Figure 1 opens to the original publications in both languages – in our case the relevant EU Regulation.

Linguee: a vast collection of human translations

Of note, Linguee is not an automatic translator like Google Translate or Microsoft’s Bing Translator. These tools, although helping you understand the gist of foreign language text, may not always use the correct term or phrase in a given context because they do not understand the subtleties of language. By contrast, what Linguee displays is human-translated entries, showing you how other people have solved a particular translation problem. Although, as with any linguistic resource or dictionary, caution is required when making your choice, Linguee is a highly valuable addition to any multilingual toolkit.

For more information, go to www.linguee.com/

Gabriele Berghammer
the text clinic
www.the-text-clinic.com

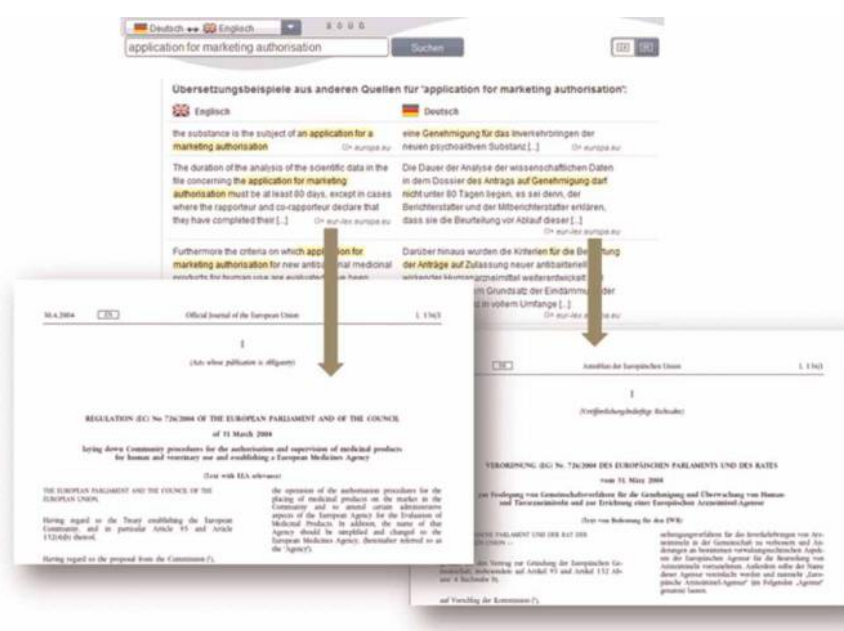


Figure 1: Linguee search result.