Creating an English-Hungarian Termbase for Laser Physics[1] The DictionELI

ÁGOTA FÓRIS
Károli Gáspár University TERMIK

ESZTER B. PAPP
Károli Gáspár University TERMIK and LEG Zrt.

KEYWORDS: DictionELI, terminological dictionary, field specific terminology policy, laser physics, English-Hungarian terminology dictionary for laser physics

1. INTRODUCTION

The paper aims to introduce a Hungarian cooperative project that had the aim of collecting English and Hungarian terminology for laser physics, and making it publicly available in a terminological database and a terminological dictionary. The collection and publication of English and Hungarian terminology for laser physics can be considered an objective in Hungarian terminology policy of laser physics. (On terminology policies see *Guidelines for Terminology Policies*, 2005; on Hungarian terminology policy, see Fóris, Sermann 2010; Fóris 2011; Fóris, Papp 2011.) Work on laser terminology is best carried out through the cooperation of field experts and linguists, more precisely with the collaboration of laser physicists and terminologists, which is appropriate to coordinate the process and publish the results.

Preparation for this pilot project began in the autumn of 2013, the project itself was launched in December 2013, and finished in April 2015. In our practice this was the first terminology project where we had the opportunity to carry out every step from start to finish according to international terminology recommendations. The aim was to collect, harmonize and publish English and Hungarian terminology for laser physics.

[1] Written as part of TÁMOP-4.1.1.C-12/1/KONV-2012-0005 project titled "Ágazati felkészítés a hazai ELI projekttel összefüggő képzési és K+F feladatokra", and also funded by the Bolyai János research grant awarded by the Hungarian Academy of Sciences.

The making of DictionELI, the terminology database on laser physics, is equally important for the development of Hungarian language for specific purposes, and for the quality and quality assurance of English to Hungarian translations. The mother tongue is a means of not only communication and identity, but also of learning. However, a community can only exercise its right to the use of the mother tongue if certain conditions are given. One of these conditions is the existence of precise, accurate and validated terminology in the given language. Terminology needs to be collected, and if it is not or only partly available, it needs to be created. A human resource condition of the collection and/or creation of terminology is the close cooperation of linguists and field experts. As for the material requirements, we need to mention an established institutional background, the availability of tools, and significant financial resources due to the time consuming nature of the process. All this effort is only fruitful however, if the terminology thus built is made publicly available, meaning the checked and validated data is published in open access databases and/or (free) online terminological dictionaries (cf. Galinski, Weissinger 2010; ISO 860:2007). The collection, creation, harmonization and publication of mother tongue terminology is the joint duty of field experts and linguists.

In the paper we present the preparation for the project, its objectives, the methods of building the terminology database and some examples from the dictionary (DictionELI).

2. BACKGROUND, PARTICIPANTS AND SOURCES OF THE PROJECT

2.1. History of Hungarian terminology for laser physics

The Hungarian terminology of laser physics, the development thereof, and as a specific example the introduction of the term *lézer* (EN 'laser'), to name the new type of light source built in 1960, into the Hungarian terminological system, was discussed by Fóris (2005: 47–56, 68–90), based on an interview with the physicist László Kozma. The first laser was built in 1960 and soon after it gained not only scientific significance, but also became a generally used tool (in medicine, in the industry, in commerce etc.). Hungarian laser research began just two years after the first US and Russian publications on the details of the first laser, and in the very same

decade research and development projects were launched to prepare application at the University of Szeged and at the KFKI of Budapest.[2]

The dissemination of research findings was done in English and Russian, while laboratory work, teaching, creating acceptance for the new scientific field and preparations for application were done in Hungarian. This made it necessary to develop the Hungarian terminology for laser physics.

The working committees of the Hungarian Academy of Sciences and the member organizations of the Federation of Technical and Scientific Societies provided professional forum and publication opportunities that made it possible to present scientific findings and introduce and discuss suggestions regarding terminology. The development of the standard on lasers by the Standards Authority offered terminology works a framework. Representatives of the relevant institutions took part in the drafting of the standard, which also contained terminological definitions. Therefore, not only the terms fixed in the standard, but also all terms of the domain were created through a professional consensus, based on the opinions of representatives of the leading institutions at the fields of research, development and application, and internationally renowned experts in science and application.

For example, they agreed to adapt the term written in English as *laser* and write it in a way that reflects its Hungarian pronunciation, namely *lézer*. The main reason behind this was that *lézer* became a term that was generally used in international scientific literature, was incorporated into many languages, and therefore can be considered an international expression. It also met most of the criteria for Hungarian terms.

From this term a large number of other terms were created observing the rules of the Hungarian language, in a way that they did not invent completely new expressions, but combined the new term with existing words and endings, etc. This way a large number of compounds were formed, e.g.: new genus terms such as *nitrogénlézer* (nitrogen laser), *folyadéklézer* (liquid laser), *CO2 lézer* (CO2 laser); *lézerkés* (laser knife), *lézernyomtató* (laser printer); *lézerterápia* (laser therapy), *lézeres anyagmegmunkálás* (laser material processing) etc.

[2] "The dictionary is dedicated to the memory of László Kozma and István Ketskeméty professors, who started laser research in Szeged." (DictionELI, frontpage, Figure 1.)

2.2. The "Superlaser" Project

The idea and possibility to compile the database of laser terminology arose with a European project, the ELI (Extreme Light Infrastructure), which is a large-scale European project. The project was launched in 2008, and requires the cooperation of a large number of research and academic institutions from 13 EU Member States. The final aim of the ELI Project is the construction of state-of-the art laser facilities, which will be used for research and application purposes. One location at which the ultrashort, high energy laser is being built is Szeged, Hungary:

> "The Extreme Light Infrastructure (ELI) project is an integral part of the European plan to build the next generation of large research facilities identified and selected by the European Strategy Forum on Research Infrastructures (ESFRI). Extreme light infrastructure will be the first infrastructure in the world to enable the investigation of the interaction between light and matter with the highest intensity, in the so-called ultra-relativistic range. It will open a doorway into new territories within physics as well as establishing such new technical developments as relativistic microelectronics and small laser particle accelerators. ELI will have a considerable impact on numerous fields of materials sciences, medicine and environment protection." (www.eli-hu.hu)

Since the first laser was built in 1960, Hungarian laser physicists have been working in the forefront of research and development, and this is why Hungary was selected as one of the locations for the international laser project (Extreme Light Infrastructure (ELI): www.eli-hu.hu). Szeged in the south of Hungary will host the attosecond research centre of ELI:

> "The main objective of ELI Attosecond Light Pulse Source (ELI-ALPS) is establishing a unique attosecond facility which provides light sources between THz (10^{12} Hz) and X-ray (10^{18}-10^{19} Hz) frequency range for developers and users in the form of ultrashort pulses with high repetition rate. Experimental projects demanding ultrahigh intensity light, like laser particle acceleration or laser generated X-ray radiation will be mainly carried out at the Beamline Facility in Prague, Czech Republic, while the photoinduced nuclear experiments will be performed at the research institute to be built in Magurele, next to Bucharest, Romania." (www.eli-hu.hu)

The location of the fourth research centre devoted to non-linear quantum electrodynamics and laboratory astrophysics has not been decided yet.

2.3. The terminology subproject

The terminology development of laser physics was carried out as part of the ELI laser project hosted by the University of Szeged, more precisely

as one task of the Támop-4.1.1.C-12/1/KONV-2012-0005 project (see http://www.u-szeged.hu/tamop411c0005-index). The aim of the terminology subproject was to compile a terminology database and an online terminological dictionary (DictionELI, http://dictioneli.stepp.hu) at the field of lasers, more precisely high energy, atto- and femtosecond pulse lasers. DictionELI contains a part of the data we collected in the database, but not all.

Cooperating partners were the Physics Department at the University of Szeged as the field experts and the Terminology Research Group at the Károli Gáspár University (KRE TERMIK) as the linguists. Six terminologists took part in the project: Ágota Fóris (KRE TERMIK), Eszter Papp (LEG Zrt. & KRE TERMIK), Eszter Sermann (SZTE & KRE TERMIK), Andrea Faludi (LEG Zrt. & KRE TERMIK), Eszter Nagy (MSZT & KRE TERMIK), Adrienn Petrányi (KRE TERMIK). Field expert reviewers were: Gábor Almási (PTE, Pécs), Miklós Erdélyi, Zoltán Horváth, Attila Kovács and Péter Makra (SZTE, Szeged), Miklós Klebniczki (KEFO, Budapest), Attila Barócsi and Pál Maák (BMGE, Budapest), Aladár Czitrovszky and Miklós Veres (SZFKI, Wigner FK, Budapest). The idea originated from Károly Osvay (SZTE, Szeged), who helped us through the project with regular advice as well. Péter Maróti (SZTE, Szeged) was the project manager, and Ferenc Havasi (SZTE, Szeged) was in charge of programming the database.

As the international project's working language is English, participants wrote all the documents prepared during research and development in English, but the documentation also had to be submitted in Hungarian. These English documents and their Hungarian translation – about laser physics, related to the ELI project – served as the basis for term extraction. The processed English and Hungarian texts contained approximately 130 000 tokens. We also extracted terms from five English and Hungarian doctoral thesis summary books. As a result we identified 5200 terminological units, and their corresponding English and Hungarian terms, along with context and other relevant information, were included in the dictionary.

Processing the texts and manually extracting terms, and several rounds of expert review was a time consuming process and therefore the project took one and a half years.

3. THE MAKING OF TERMINOLOGY DICTIONARY OF LASER PHYSICS (DictionELI)

3.1. Corpus and working methods

The aim of the terminology subproject was to compile a terminology database and an online terminological dictionary (DictionELI) in the field of lasers, more precisely high energy, atto- and femtosecond pulse lasers.

The first step was to set the principles and methodology of building a terminology database. Another initial thing to do was to implement the methods of terminology to the field of the terminology of laser physics, and to optimize applied terminology methods and make them more effective.

The bilingual (English and Hungarian) texts that were written in connection with the ELI project were handed over to terminologists in an electronic format; these texts made up the research corpus, and the bilingual term extraction was carried out from them. The manually extracted data were stored in an Excel table. This process had the following steps:

1. extracting source term candidate (indicating the source),
2. selecting the definition (if found in the text),
3. extracting target term candidate(s) (indicating the source),
4. when having several target language alternatives, checking in reference works (and with laser physicists),
5. checking and harmonizing spelling,
6. review by field experts (laser physicists).

Future terminology standardization will be significantly helped if computer assisted translation (CAT) tools are also available, which could incorporate former translations in their translation memory, and the Excel table could be easily converted to a terminology database. Such a translation memory and termbase could help minimize translation costs and reduce errors, and a consistent and harmonized terminology use enhances the coherence of the created text.

The most important objectives of building the terminology database were:

– defining necessary data fields,
– filling up the database with the collected data,
– making the database available for translations (as it can be incorporated into a CAT tool).

The database is never considered finished, it is continuously expanded with new translations.

3.2. A few examples for the laser terminology database

In the following section we are going to illustrate the data fields of the Excel table containing the laser terminology through a few examples (see Table 1): the English source term (EN term) / its part of speech (part of speech) / what it is an abbreviation of (abbreviation of) / context of the English term (EN context) / the Hungarian term (HU term) / Hungarian synonyms (HU synonym) / the category of the Hungarian term (HU category) /source of the Hungarian term (HU source) / context of the Hungarian term (HU context) / other information (other).

On the user surface of the DictionELI not all above information is visible.

Take an example of the English verb *absorb*, and its Hungarian equivalent *elnyel*. We indicated both the English and the Hungarian context in the table, its category is "jóváhagyva", meaning 'approved', and you can also see the source (which is important to note because it makes it trackable and easy to check).

The English abbreviation *AFA-LIFT* stands for "absorbing film assisted laser induced forward transfer", in Hungarian the English abbreviation is used: *AFA-LIFT*.

The English noun *amplified spontaneous emission* means in Hungarian *erősített spontán emisszió*. We indicated the context, the category, sources and abbreviation (ASE) here as well.

The English noun *chirped mirror* appeared to have two Hungarian equivalents in two different source texts: *fázismoduláló tükör* and *csörpölt tükör*. The first Hungarian term reflects the function of the mirror, while the second one includes the Hungarian pronunciation of the English word "chirped".

The English term *clean room* appeared in our source texts in three different spelling variants *clean room*, *clean-room*, *cleanroom*, and we found two Hungarian equivalents: *tisztatér* and *sterilszoba* (see Table 1).

We recorded terms having different spelling or having several Hungarian equivalents as separate entries (as they are homographs), and we leave it to the experts to decide whether they refer to different or the same concepts.

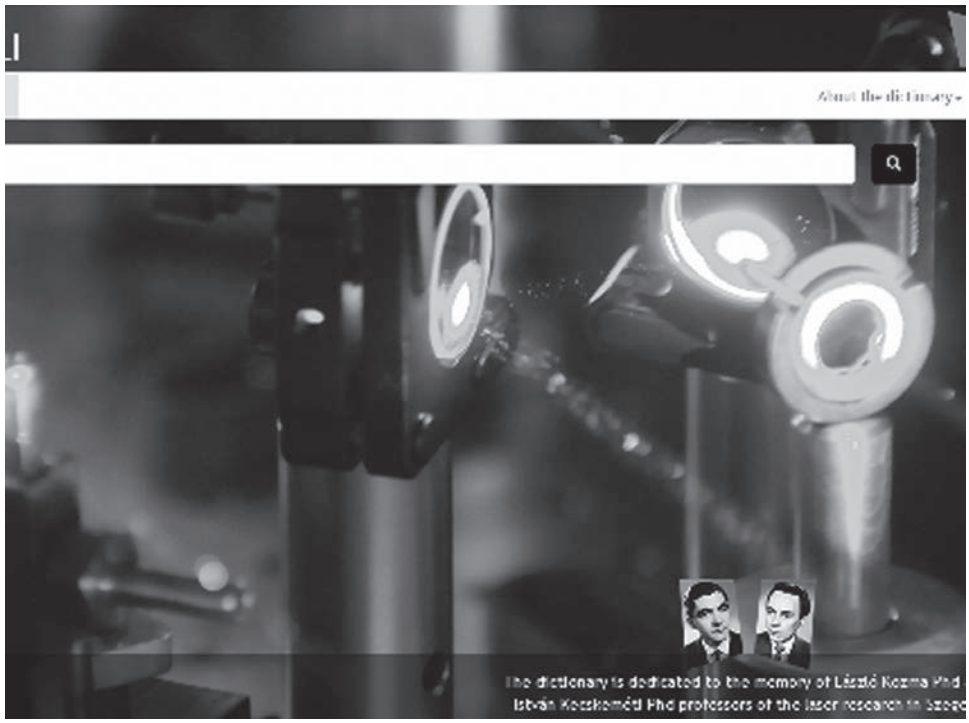
Table 1. Examples for the laser physics terminology database

EN term	part of speech	abbreviation of	EN context	HU term	HU synonym	HU category	HU source	HU context	other
absorb	verb		Multiple passes of the pump light through the thin disk are realized by appropriate pump optics in order to absorb most (typically > 90%) of the pump radiation.	elnyel		jóváhagyva	4_le_3.doc	A pumpa megfelelő optikája segítségével elérik, hogy a pumpáló fény többször is áthaladjon a korongon és a pumpáló sugárzás nagy része (jellemzően több mint 90%-a) elnyelődjön.	
AFA-LIFT	abbreviation	absorbing film assisted laser induced forward transfer	In the Institute of Physics USZ a new laser-based method for controlled transfer of biological samples is developed, namely the absorbing film assisted laser induced forward transfer (AFA-LIFT).	AFA-LIFT	elnyelő filmmel segített, lézerrel indukált, előre irányuló átvitel	lektorált	ELI 84_2014 - 02 6_ 6_J.Bohus	Az USZ Fizikai Intézetében a biológiai minták szabályozott átvitelének új, lézeralapú módszerét, nevezetesen az elnyelő filmmel segített, lézerrel indukált, előre irányuló átvitelt (AFA-LIFT) fejlesztik ki.	

EN term	part of speech	abbreviation of	EN context	HU term	HU synonym	HU category	HU source	HU context	other
amplified spontaneous emission	noun		The limitation due to the amplification process in laser amplification media is caused by spontaneous emission occurring at the beginning of the amplifier chain that is then amplified along its length and is commonly referred to as amplified spontaneous emission (ASE).	erősített spontán emisszió		jóváhagyva	4_le_1.doc	A lézer erősítő közegében lejátszódó erősítési folyamat miatti korlátozást az erősítő lánc kezdetén jelentkező spontán emisszió okozza, amely aztán felerősödik annak hossza mentén, és amelyet általában erősített spontán emissziónak (ASE) neveznek.	ASE
chirped mirror	noun		Pulse intensity considering additional dispersion introduced using bulk glass or chirped mirrors to prestretch the pulse duration on the first grating around 1-ps, which limits nonlinear effects and damage to optics.	fázismoduláló tükör		jóváhagyva	3_SYLOS_01_le61	Impulzusintenzítás, figyelembe véve az impulzushossznak az első rácson történő, – a nemlineáris hatásokat és az optika károsodását csökkentő – kb. 1 pikoszekundumos előnyújtásához szükséges tömbi üveg vagy fázismoduláló tükrök által keltett többletszórását.	
chirped mirror	noun		The pulse can be stretched by material and recompressed by chirped mirrors.	csörpölt tükör		jóváhagyva	4_le_1.doc	A lézerimpulzus megnyújtható anyag segítségével, és újra-összenyomható csörpölt tükrökkel.	

EN term	part of speech	abbreviation of	EN context	HU term	HU synonym	HU category	HU source	HU context	other
clean room	noun		The procedure requires an appropriate technology and equipment, such as a clean room, an interferometer and various handling devices for the disks.	tisztatér		jóváhagyva	4_le_3.doc	Az eljáráshoz megfelelő technológiára és felszereltségre van szükség, például egy tisztatérre, egy interferométerre és különböző fogó eszközökre a korongokhoz.	
clean-room	noun		In order to be independent of external deliverers and to have the highest degree of flexibility during experimental campaigns, access to a clean-room area equipped with at least micro machining with and without fs-lasers, UV- (and possibly electron-) lithography, and spin coating systems is required.	sterilszoba		jóváhagyva	ELI 84_2014 - 03.6_11_ER_IB_ER	Ahhoz, hogy a külső beszállítóktól függetlenek legyünk és a lehető legnagyobb flexibilitással rendelkezzünk a kísérletek során, elengedhetetlen egy hozzáférhető sterilszoba, amelynek legalább a következő berendezéseket kell tartalmaznia: femtoszekundumos lézerrel és anélkül működő mikromegmunkáló, UV- (és lehetőleg elektron-) litográfia, valamint forgótárcsás rétegképző (spin coating) rendszer.	
cleanroom	noun		The laser should be in a class ISO 7 cleanroom, the spatial filters can operate with their own dry vacuum pumps.	tisztatér		jóváhagyva	4_le_4.doc	A lézert ISO 7 osztályú tisztérben kell elhelyezni, a térszűrők a saját olajmentes vákuumszivattyúikkal üzemelhetnek.	

Figure 1. DictionELI, frontpage



4. CONCLUSIONS

We consider the making of DictionELI a significant achievement for several reasons. 1) This is the first terminology project in our practice that has provided us the opportunity to carry out every step from start to finish according to the international terminology recommendations. Experts – laser physicists – provided us with all required texts, allowed us to convince them of the correctness and necessity of the working methods, and carried out the review and validation quickly and precisely. Cooperation between linguists and physicists was an example to follow. 2) At the time of publication the online terminology database contained over 5000 entries, and it can be further expanded. It is an open access database, which means it is freely accessible from anywhere in the world. This makes it possible for physicists and university or doctoral students to use and check unified Hungarian terms. Therefore DictionELI can also be considered an achievement in Hungarian terminology policy of laser physics. 3) When doing specialized translation, the existence of a

terminology database is a prerequisite of the quality control of translations. By creating DictionELI and making it publicly available, translators and translator agencies have access to the English and Hungarian terminology of high energy pulse lasers, and as a result creating terminologically sound, coherent texts becomes easier, quicker and cheaper, which reduces translation costs and helps avoid translation errors.

We are planning to keep expanding the dictionary in the English–Hungarian language pair, as the building of the laser facility and the actual research carried out will bring about a large number of written texts, which can serve as basis for further term extraction.

We also hope that this dictionary could be a starting point for laser terminology dictionaries in other language pairs, and eventually a multilingual (e.g. English, Hungarian, Czech, French, Romanian, maybe Lithuanian, etc.) laser terminology dictionary.

SOURCES

DictionELI – http://dictioneli.stepp.hu [available from 1st June 2015].

Extreme Light Infrastructure (ELI) – www.eli-hu.hu [accessed May 2015].

Guidelines for Terminology Policies. Formulating and implementing terminology policy in language communities, Paris: UNESCO, 2005 – http://unesdoc.unesco.org/images/0014/001407/140765e.pdf [accessed May 2015].

ISO 860:2007 *Terminology work – Harmonization of concepts and terms*.

TÁMOP project – http://www.u-szeged.hu/tamop411c0005-index [accessed May 2015].

REFERENCES

Fóris Á. 2005: *Hat terminológia lecke* [Six Lectures on Terminology.], Pécs: Lexikográfia Kiadó.

Fóris Á. 2011: A magyar terminológia-politika elvei, a gyakorlat és az elmélet viszonya [The principles of Hungarian terminology policy, the relation of theory and practice.] – Hires-László K., Karmacsi Z., Márku A. (eds.). *Nyelvi mítoszok, ideológiák, nyelvpolitika és nyelvi emberi jogok Közép-Európában elméletben és gyakorlatban. A 16. Élőnyelvi Konferencia előadásai*, Budapest–Beregszász: Tinta Könyvkiadó – II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola Hodinka Antal Intézete, 421–428.

Fóris Á., B. Papp E. 2011: A terminológiai szabványosítás és harmonizáció a nyelvi jogok érvényesítésének szolgálatában. [Terminology standardization and harmonization serving the vindication of linguistic rights.] – Hires-László K., Karmacsi Z., Márku A. (eds.). *Nyelvi mítoszok, ideológiák, nyelvpolitika és nyelvi emberi jogok Közép-Európában elméletben és gyakorlatban. A 16. Élőnyelvi Konferencia előadásai*, Budapest–Beregszász: Tinta Könyvkiadó – II. Rákóczi Ferenc Kárpátaljai Magyar Főiskola Hodinka Antal Intézete, 429–433.

Fóris Á., Sermann E. 2010: A terminológiai szabványosítás és a terminológiai harmonizáció. – *Magyar Terminológia* 3/1, 41–54.

Galinski Ch., Weissinger R. 2010: Terminológiai szabványosítás és fordítási szabványok. – *Magyar Terminológia* 3/1, 8–20. [Translated by Eszter B. Papp: Galinski Ch., Weissinger R. 2008: Terminology standardization and translation standards. – XVIIIth FIT World Congress / XVIIIe Congrès mondial de la FIT / Daishiba-jie Shijie Fanyi Dahui. *Translation and Cultural Diversity / Traduction et diversité culturelle / Fanyi yu duoyuan-wenhua.* Proceedings / Actes / Lunwenji, Shanghai: Foreign Languages Press (CD-ROM). 11.]

ANGLŲ–VENGRŲ KALBŲ LAZERIŲ FIZIKOS TERMINŲ BAZĖS DictionELI KŪRIMAS

Straipsnyje aprašomas terminologijos darbas renkant vengrišką lazerių fizikos terminiją ir pristatoma duomenų bazė DictionELI (http://dictioneli.stepp.hu), kurioje skelbiama sukaupta informacija. Lazerių terminijos rinkimas ir skelbimas gali būti laikomas Vengrijos lazerių fizikos terminologijos politikos tikslu. Šis darbas yra bandomasis projektas, kurį įgyvendinant kalbininkai ir srities specialistai turėjo galimybę bendradarbiauti derindami vienos iš fizikos sričių – didelės energijos, ato- ir femtosekundinių impulsinių lazerių – terminiją. Šis taikomasis tyrimas gali būti naudojamas kaip atskirų sričių terminijos tyrimo ir plėtros Vengrijoje modelis.

Gauta 2015-06-15

Ágota Fóris
Károli Gáspár University TERMIK
1146 Budapest, Dózsa György út 25–27. (Hungary)
E-mail: foris.agota@gmail.com

Eszter B. Papp
Károli Gáspár University TERMIK and LEG Zrt.
1146 Budapest, Dózsa György út 25–27. (Hungary)
E-mail: eszter.papp@gmail.com