Inkaliisa Vihonen-Peuranen

The Language Industry Platform LIND and national market surveys – why data matters

Abstract (Finnish)

Kieliteollisuus ja sen osa-alueet kääntämisestä ja tulkkauksesta aina kielikonsultointiin ovat olleet Euroopan komission kääntämisen pääosaston (DGT – Directorate-General for Translation) kiinnostuksen kohteena jo yli vuosikymmenen, aina vuonna 2009 julkaistusta tutkimuksesta kieliteollisuuden koosta Euroopassa. Monikielinen Eurooppa on synnyttänyt palveluita, jotka mahdollistavat viestien välittämisen vastaanottajan omalla äidinkielellä. Tämä on niin yhteiskunnan kuin liiketoiminnan kannalta välttämätöntä ja järkevää.

Kieliteollisuutta on kuitenkin varsin vaikeaa määritellä, monet sen toiminnoista ovat näkymättömissä eri prosesseiden syövereissä, joskus jopa täysin irrallisena niistä. Kielten elinvoimaisuuden ja dynamiikan kannalta on kuitenkin olennaista tuoda esiin näiden palveluiden laajuus ja merkitys nykypäivän Euroopalle. Euroopan Unionin (EU) tasolla onkin luotu erinäisiä hankkeita, joiden tarkoituksena on luoda näkyvyyttä erilaisille palveluille, joita kieliteollisuus tuottaa.

Olennainen osa näkyvyyttä on kieliteollisuuden esiintyminen virallisissa tilastoissa. DGT on tehnyt yhteistyötä alan asiantuntijoiden kanssa eritelläkseen luokkia, joihin kielipalvelut kuuluvat muun muassa Euroopan laajuisissa EU:n tilastotoimisto Eurostatin tilastoissa. Työ on hidasta, muutoksien nopeutta mitataan vuosissa tai joissain tapauksissa vuosikymmenissä. Toisaalta politiikan tai rakenteellisella tasolla tehdyt muutokset ovat pitkällä aikavälillä niitä tehokkaimpia näkyvyyden tuojia. Tätä pitkän aikavälin työtä tukemaan on järjestetty useita Euroopan laajuisia ja kansallisia kyselyitä, jotka puhuvat omaa kieltään monimuotoisesta mutta toisaalta pirstaloituneesta alasta. Tarvitaankin kaikkien alan toimijoiden yhteistyötä kokoomaan yhteen tärkeän, mutta hajanaisen kielisektorin tieto ja osaaminen. Vain yhdessä ja kaikkien osa-alojen yhteistyöllä saavutetaan näkyviä ja kestäviä tuloksia.

1. Some milestones in the quest for visibility

The language industry and its various sub-sectors have been a topic of interest and research at DGT – Directorate-General for Translation – since 2009, when the first mapping exercise, *The study on the size of the language industry*, was published. With its preliminary estimates and definition of sub-sectors, it marked a first step in the quest of for visibility and recognition of a rapidly developing industry. As well as presenting the situation back in 2009, it gave estimates of how the industry would develop in the future.

During a time when there were far fewer multilingual global services like streaming services on multilingual platforms, the estimated value of the language industry was roughly €8.6 billion. It was also expected that the size of the industry would double in the following eight years and be around €16.8 billion in 2015.

Today, a good decade later, the figures seem quite realistic even though there is still no exact figure on the value of the European language industry in total. However, moderate estimates vary between €15-20 billion. Companies involved in the language industry also report on growth and optimism about language market developments. In a word, the volume of multilingual services has increased, much for the benefit of citizens of multilingual Europe.

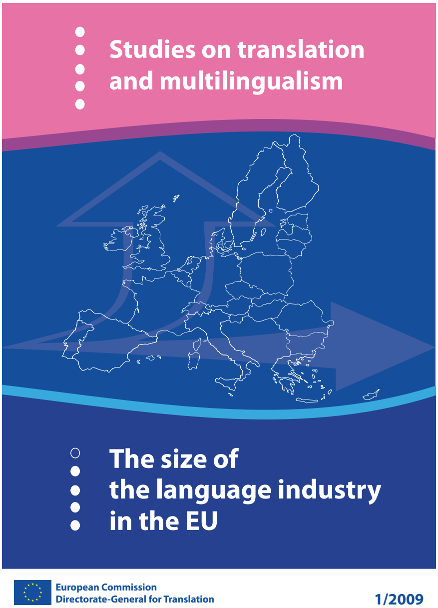

Fig. 1: Study on *The size of the language industry in the EU* published in 2009

2. What is the language industry composed of?

One of the other, more lasting results of the study was the classification of the language industry into 8 sub-sectors as follows:

- Translation,
- Interpretation,
- Software localisation and website globalisation,
- Language technology tools development,
- Language teaching,

- Subtitling and dubbing,
- Organisation of international conferences,
- Consultancy in linguistic issues.

Again, the idea was to demonstrate that there are many versatile services enabling multilingual content creation for citizens and businesses besides translation and interpretation, the most well-known and visible sub-sectors of all.

Today, the situation of the industry seems even more versatile; with the involvement of *Artificial Intelligence* (AI) and *big data*, many new services have emerged, creating more products to cater for linguistic needs. Examples of such new services are *audio description*, *text-to-speech services* and *transcreation*.

3. Statistics and more statistics

From the start, it was clear that some structural measures were needed to enhance the visibility of the language sector. From an outside perspective, it is hard to recognise something that is not visible and not easily quantifiable.

In Europe, we are in a lucky situation where we do get free and comprehensive data on, for example, European business sectors. However, the language industry, for historical reasons, has not been represented separately but together with other sectors, such as secretarial services. Therefore, it has been hard to demonstrate the size and development of the sector.

To this end, DGT has worked closely with the Statistical Office of the European Union *Eurostat*. During the last review of the business activity classification NACE (Statistical Classification of Economic Activities in the European Community) that ended in 2008, the sub-sectors of translation and interpretation were already separated from secretarial services, yielding data that is much more usable and reliable. In the latest review launched in the summer of 2019, we proposed to separate the categories of translation and interpretation, currently intertwined, as they usually represent different professions and skillsets. In addition, separate categories were proposed for activities in machine translation and speech recognition.

As another follow-up to the initial study, DGT established the *Commission Expert Group and Platform for the Language Industry, LIND*. In this context, DGT has worked together with LIND experts, all recognised experts of the language profession, to enhance existing and future tools to map industry sectors, for example, in ESCO (European Skills, Competences, Qualifications and Occupations), which currently covers European job titles in 24 languages. There the 'new' language industry occupations include *lawyer linguist, localiser* and *language engineer*. These new occupations also link in to the EU-wide *Eures* database, where job seekers can find work across Europe. Hopefully, this will also increase the visibility and demonstrate the potential that the language industry has to offer to them.

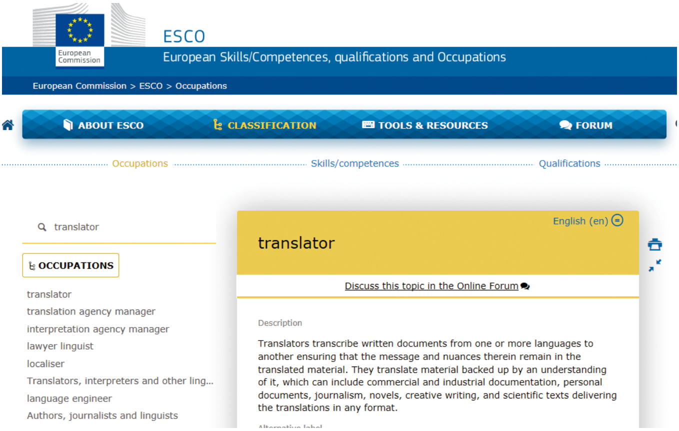

Fig. 2: Overview of language professionals in the ESCO classification

Besides these activities to enhance data acquisition relating to the industry which DGT is itself an essential part of, we have cooperated in the voluntary collection of data to track and identify industry trends, which will help us to monitor and develop our future activities. These initiatives include the *European Language Industry Survey* and various national surveys carried out in EU member states using the European survey as a model.

4. What next?

Now, ten years later, it has become clear that the positive image the pioneering study painted of the language industry back in 2009 was quite realistic or even slightly cautious, as the industry has continued to grow due to several, sometimes converging, factors such as globalisation, new market regulations and diversifying customer needs.

Nevertheless, it is still a relatively small and fragmented sector due to the innate nature of languages. On a positive note, the diversity of the language industry makes it also dynamic and innovative. In the age of automation, it is highly unlikely that this diversity and complexity can be mastered by automated tools any time soon.

The fragmentation of the language industry also calls for close cooperation between small sub-sectors and professions across languages and professional borders. Through cooperation we can make our case for language services being at the heart of the European project. It is only by joining forces as well as sharing the knowledge and expertise that we all represent that we will gain visible and sustainable results for the language industry as a whole.

Bibliographical information

This is a preliminary version of the text which will be published in the book:

Tender, Tõnu/Eichinger, Ludwig M. (eds.): Language and economy. Language industries in a multilingual Europe. Contributions to the EFNIL Conference 2019 in Tallinn. Budapest: Research Institute for Linguistics, Hungarian Academy of Sciences.

The electronic PDF version of the text is accessible through the EFNIL website at:
http://www.efnil.org