



Erasmus+ – Cooperation partnerships in the field of Higher Education (KA220-HED)

2021-1-TR01-KA220-HED-000027580
EDUSIGN & SIGNEDU



JOINT VENTURE OF CURRICULUM STUDIES AND ARTIFICIAL INTELLIGENCE IN SIGN LANGUAGE

Project Result 2	Artificial Intelligence Based Sign Language Interpreting Infrastructure in the Context of Finnish Sign Language
-------------------------	--

Work Package 2 (R2) focused on developing an algorithm based on artificial intelligence (AI) that was initially produced to recognize written Turkish language and to produce raw translations in Turkish Sign Language. This algorithm was to be accommodated into written Finnish and Finnish Sign Language (FinSL). It was hoped that this application developed during the three years of this Erasmus project would bring us an exceptional occasion to study the possibilities of machine translation between Finnish and FinSL, since the subject is a hot topic in the field of sign language research.

Initially, in the Grant Agreement, 120 days (and 25 680 €) was reserved for Humak for this subproject, while for developing the AI algorithm 96 work days (and 7104 €, *sic!*) were granted for the Istinye University in order to create an algorithm able to create equivalent phrasal structures between written Finnish and FinSL. It was more than evident that only the initial part of the development project could be covered with that amount. Thus, during the first year of the project Humak's team gathered for the Turkish partners from openly available materials (as the Language Bank of Finland) all relevant information on the algorithms developed in Finland for automatic recognition of written Finnish and on the digital databases on FinSL.

In Autumn 2023, with the help of the re-funding arrangements, we were finally able to provide the algorithm developers the first collection of 100 random, simple-structured sentences in FinSL with their phrasal equivalents in written Finnish. With these sentence examples, collected from an online digital dictionary (suvi.viittomat.net), the AI experts were able to assess the complexity of their future work. Humak's team translated the sentences into English; tagged them with glosses (verbal equivalents reflecting the word order in FinSL) and with linguistic symbols marking the generic phrasal structure. This material was produced on an Excel spreadsheet, each record for one sentence on its own line. – Cf. Fig. 2–3.)

With the help of this list of 100 phrasal examples, we were able to discuss face-to-face, at the TPM meeting in Istanbul (August 2023) various problems related to the algorithm developers' coding model (Fig. 1). We also found that Humak's team is not able to answer with certainty to all the questions raised by the Turkish coders, in particular those related to sign order and sign choice, because the grammatical description of Finnish sign language is still in progress.

This is why we decided to focus on linguistically more studied phrasal structures (e. g. interrogative and negative phrases) and concentrate on extracting only those from the digital online dictionary on FinSL. During the Autumn 2023 and Spring 2024, we had extracted and analysed a total of 160 sentences containing a question sign (Fig. 2) and ca. 125 sentences without a question sign (so-called YES-NO-questions, Fig. 3) and, 100 negative sentences, hoping that the algorithm developers would have project time to develop.

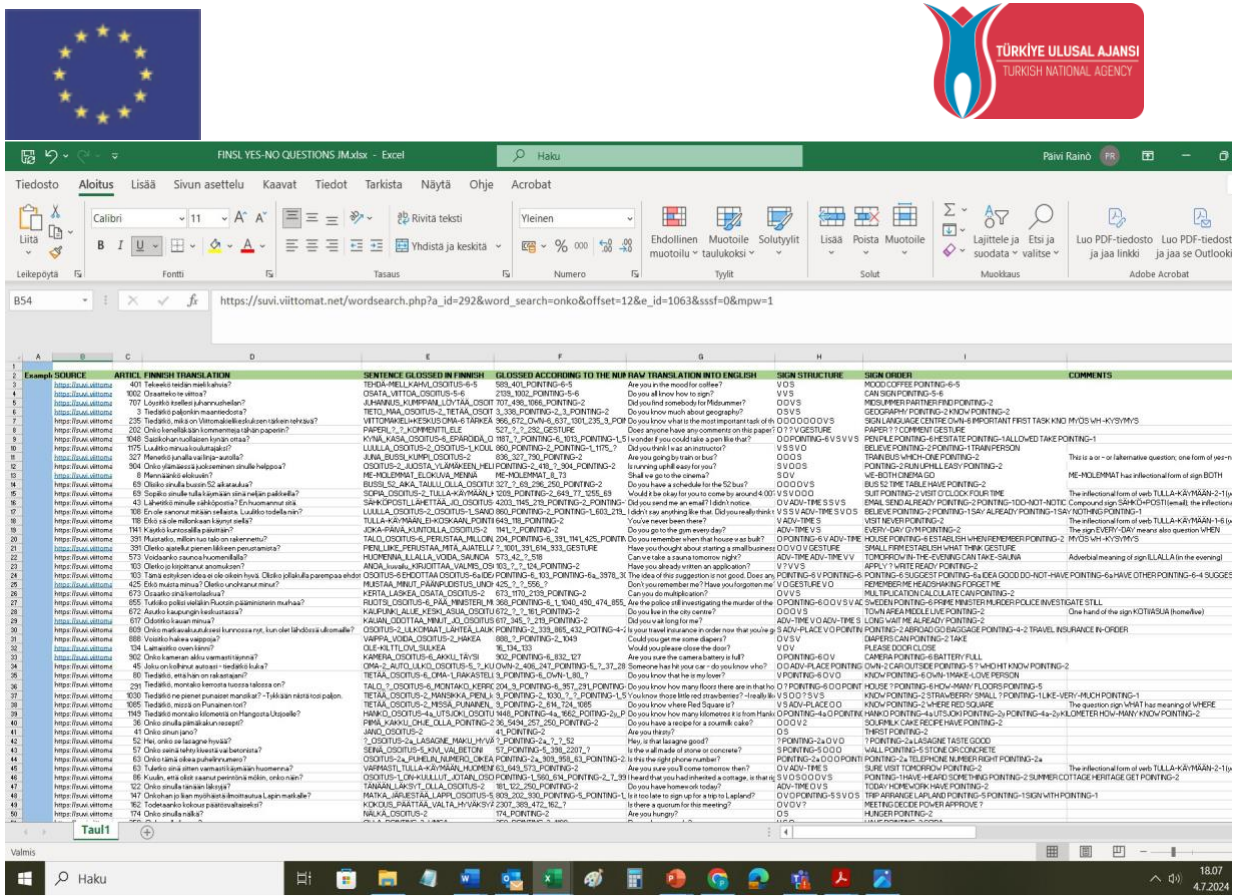


Fig. 3. Examples on analysed YES-NO -question phrases in FinSL

During the project, we had several interesting online discussions with the algorithm development team about the recurrent features they had noticed in FinSL. Interestingly, these features are familiar to the researchers of FinSL, but most of them have not yet been described in the existing FinSL grammars, mainly focusing on the elementary level of proficiency.

The discussions also brought to surface several other “problematic” features [found not only in FinSL but in all signed languages], of which the following two are most challenging for all algorithm developers: 1) The low degree of conventionalisation of signs and, 2) high context dependency of the signed phrasal structures. - Nevertheless, these were already known beforehand, and this information was also brought to everyone's attention by Humak's team when the project was planned. These features were described in the following way in the article "Is There Any Hope for Developing Automated Translation" (Cf. Jantunen, Rousi, Raino & al. 2021), published before the start of this Erasmus + project:

In the investigation of corpora of Western sign languages, it has been estimated - - that approximately 30 percent of sign tokens represent *depictive content that cannot be unambiguously translated even in the traditional sense* - - . The percentage in itself reveals the situation. However, this is made even more confusing by an estimation that the group of 70 percent of non-depictive signs comprises only approximately 3000 lexeme types. If we consider all of this, from the point of view of the current Corpus FinSL data - - for instance, the end result is that we can perhaps fairly easily translate (in the traditional sense) approximately 75500 sign tokens with the help of 3000 lexeme types. Yet, *more challenges will undoubtedly be encountered regarding the estimated 33500 sign tokens that cannot be reduced into lexeme types. This is due to the unconventionalality - or context dependence - of their form and semantic value.* Another issue that needs to be considered is that depictive meanings are expressed not only with manual signs (i.e. with hands) but also non-manually, with facial expressions, bodily postures and movements - - . (Jantunen & al. 2021; italics added by the author of this report).

(It is also worth mentioning here that an international EU-funded project, that promised to solve translation problems between several European written and signed languages and which received 100 times more funding than ours, faced the same problems. The Sign-On project, completed in 2023, was granted with 5.6 million euros from the Horizon 2020 RIA programme (SignON - Sign Language Translation Mobile Application and Open Communications Framework; <https://signon-project.eu/publications/other-publications/>), but resulted in most simple Avatar translation samples.)

As mentioned above, our own material contained and brought to the surface quite a many of aforementioned unconventional elements that were marked in our examples as “gestures”. The frequent use of (half-



conventionalized) gestures make it extremely difficult to create a schematic, mathematical model (cf. Fig. 1) of any signed phrase, which in turn prevents the production of a ubiquitous algorithm that would work everywhere, and even in most restricted frame work (e. g. “interrogative phrases”). To our surprise, however, at the end of our project the algorithm succeeded to produce fairly idiomatic descriptions (i. e. theoretical phrasal structure codified with in Finnish glosses) leading to 50 understandable signed phrases in Finnish Sign Language (Fig.4).

RANDOM EXAMPLES IN FINNISH	AI Translation Results into FinSL	TRANSL. IN ENGLISH
Mikä sinun puhelinnumerosi on?	PUHELIN_NUMERO_OSOITUS-2_MITÄ KYSYMYSMERKKI	What is your phone number?
Mikä tämä on?	MIKÄ_TÄMÄ_ELE KYSYMYSMERKKI	What is this?
Mikä sinun nimesi on?	NIMI_OSOITUS-2 KYSYMYSMERKKI	What is your name?
Mikä päivä tänään on?	PÄIVÄ_TÄNÄÄN_MITÄ_ELE KYSYMYSMERKKI	What day is it today?
Mikä hän on ammatiltaan?	AMMATTI_MITÄ_OSOITUS-1 KYSYMYSMERKKI	What is his profession?
Mitä sinä katsot?	MITÄ_KATSOA_OSOITUS-2 KYSYMYSMERKKI	What are you watching?
Mitä sinä opiskelet?	MITÄ_OPISKELLA_OSOITUS-2 KYSYMYSMERKKI	What are you eating?
Mitä kieliä sinä osaat?	MITÄ_KIELI_OSATA_OSOITUS-2 KYSYMYSMERKKI	What are you studying?
Mitä sinä teet illalla?	ILTA_MITÄ_TEHDÄ_OSOITUS-2 KYSYMYSMERKKI	What languages do you speak?
Mitä kuuluu?	MITÄ-KUULUA	What do you do in the evening?
Mitä sinä syöt?	MITÄ_SYÖDÄ_OSOITUS-2 KYSYMYSMERKKI	How are you?
Mitä opiskelijat tekevät?	OPISKELIJA_MITÄ_TEHDÄ KYSYMYSMERKKI	What do students do?
Mitä teet työksesi?	MITÄ_TEHDÄ_TYÖ_ELE KYSYMYSMERKKI	What do you do for work?
Mitä kieltä sinä puhut?	MITÄ_KIELI_PUHUA_OSOITUS-2 KYSYMYSMERKKI	What language do you speak?
Kuinka voitte?	KUINKA_VOIDA_OSOITUS-2 KYSYMYSMERKKI	How are you?
Kuinka paljon tuo auto maksaa?	AUTO_OSOITUS-6_MAKSAA_MONTAKO KYSYMYSMERKKI	How much does that car cost?
Kuinka monta ammattia sinulla on?	KUINKA_AMMATTI_OSOITUS-2_MONTAKO KYSYMYSMERKKI	How many professions do you have?
Kuinka pääsit sisään?	KUINKA_PÄÄSTÄ_SISÄÄN KYSYMYSMERKKI	How did you get in?
Kuinka jaksat valvoa noin pitkään?	KUINKA_OSOITUS-2_VALVOA_PITKÄ-AIKA KYSYMYSMERKKI	How can you stay up so late?
Millä bussilla menet töihin?	MILLÄ_BUSSI_MENNÄ_TÖIHIN_OSOITUS-2 KYSYMYSMERKKI	What bus do you take to work?
Millä tavalla sait kaapin ovet auki?	MITEN_OSOITUS-2_SAADA_KAAPPI_OVI_AUKI KYSYMYSMERKKI	How did you get the cupboard doors open?
Missä sinä olit?	MISSÄ_OLI_OSOITUS-2 KYSYMYSMERKKI	Where were you?
Missä kaupungissa on kuurojen koulu?	KUURON_KOULU_MISSÄ_ELE KYSYMYSMERKKI	What city has a school for the deaf?

Figure 4. Examples of phrasal structures in FinSL proposed by an AI based algorithm (in the middle column), translated from written Finnish into FinSL. For comparison, the right column shows the AI-based phrase translations from Finnish to English.

In addition, in the final phase of the project the Turkish R2-team succeeded in producing a dozen of phrasal examples using their algorithm and open resource data available at the online dictionary of FinSL (<https://suvi.viittomat.net>). The team was obliged to use a limited set of video reproductions where the quality and contents are most consistent, showing an astonishingly natural flow of sign language. The examples can be viewed on the project's website presenting Finnish partner in the EduSign project (<https://edusignedu.com/fi/finsl/>). – These examples encourage us to look for further funding opportunities for a project to develop an AI-based translation program for Finnish sign language for both sign language education and interpreter training.

We would also like to mention, as another example, how one of the questions posed by the Turkish team (Fig. 5), that led Humak's team to new information on FinSL: One of their questions on a frequent gestural sign in our analyses, brought back to surface an utmost regular element in FinSL that, should, without hesitation be given the role of a question morpheme and not a mere “gesture”. The existence of this element was already established in the early 20th century by the father of FinSL linguistics (the late D. F. Hirn, cf. Fig.5).

The fact that it is also a recurrent (but unrecognized) element in the contemporary FinSL means that modern researchers may have lost their sensitivity and maybe also their interest in focusing on the smallest elements in FinSL while observing sign language in a more multidisciplinary context, such as effects of sign language use on the brains, or the language policy status of signed languages in different countries.



*)'GESTURAL SIGN' Juha is referring to this. Why it so?

- KUUME_ELE_614_ELE KYSYMYSMERKKI
- TEE_PUSSI_MISSÄ_ELE KYSYMYSMERKKI"

"Because - and even though this very sign (see below) has existed in signed questions as "a question marker" since the very beginning of the history of our sign language (I've just checked the historical sign language dictionaries reflecting the signing in 19th century in Sweden and Finland) - it has not been "accepted" as "a true sign" in our more modern FinSL dictionaries. There is no entry for it! This proves that we are, in a way, prisoners of our glossing tradition where "sign is not a true sign if it has not been given a Finnish word-equivalent"

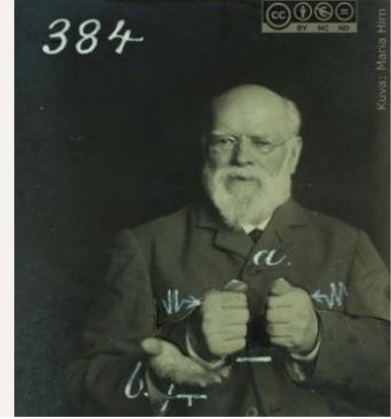


Figure 5. An example of how the reflections of the algorithm developers will lead to grammatical refinements in the linguistic description of FinSL.

However, many of those questions and discussions raised were and are still relevant and will lead to further reflections in the field of FinSL research and maybe to a new grammatical description of FinSL – e. g. a reference grammar. During the Spring 2024, this issue has already been addressed in discussions between Humak's team and developers of the online sign language dictionary (Suvi) working at the Finnish Association of the Deaf.

Link of the PR2 outcome: <https://edusignedu.com/fi/finsl/>

Video Link: https://drive.google.com/file/d/136NYJkDcS9iPBsMwi0t789cx1PTA-fwf/view?usp=drive_link