Investigating the Medical Coverage of a Translation System into Pictographs for Patients with an Intellectual Disability

Magali Norré
CENTAL, UCLouvain, Belgium
FTI, UNIGE, Switzerland

Vincent Vandeghinste
Instituut voor de Nederlandse Taal
The Netherlands

Thomas François
CENTAL, UCLouvain, Belgium

Pierrette Bouillon
FTI, UNIGE, Switzerland

Abstract

Communication between physician and patients can lead to misunderstandings, especially for disabled people. An automatic system that translates natural language into a pictographic language is one of the solutions that could help to overcome this issue. In this preliminary study, we present the French version of a translation system using the Arasaac pictographs and we investigate the strategies used by speech therapists to translate into pictographs. We also evaluate the medical coverage of this tool for translating physician questions and patient instructions.

1 Introduction

Many people around the world face difficulties to communicate through speech. To overcome this challenge, disabled people, including persons with an Intellectual Disability (ID), resort to Augmentative and Alternative Communication (AAC) systems in different forms: objects, visual aids on paper or technologies (Beukelman and Mirenda, 1998). Both text and pictographs can be used in AAC for enhancing the communication and the social inclusion of individuals with ID.

Images are already used in various medical contexts to increase access to information and communication for all, e.g. in pharmacology on drug leaflets for improving the health literacy or in hospitals for facilitating the medical tourism (Nandy, 2019). In emergency settings, there are tools such as interpreters and communication technologies for allophones patients (Janakiram et al., 2021). However, they have drawbacks and are not designed for people with disabilities. Recent research focuses on medical applications with images for disabled patients, but does not use NLP techniques (Norré et al., 2021b), such as My Symptoms Translator (Alvarez, 2014) or the system of Wołk et al. (2017).

This paper focuses on the French version of a translation system using the Arasaac pictographs[1] (Norré et al., 2021b) and makes two contributions: investigates the strategies used by speech therapists for translating medical sentences into pictographs and evaluates the lexical coverage of the system. Section 2 summarizes related work on the translation systems with images and their evaluation. We describe our system in Section 3. Section 4 investigates translation strategies. Finally, we evaluate the medical coverage of our system in Section 5.

2 Related Work

Pictographs represent one or several concepts: object, verb, feeling, grammatical word, etc. There are several pictograph sets available, such as Sclera, Beta or Arasaac, which are specifically created for people with various disabilities. They can be seen as simplified languages (Sevens et al., 2017). Therefore, they have been used within AAC systems, but many relate to daily communication.

As regards the medical language, Glyph (Bui et al., 2012) – which is not an AAC application – automatically translates patient instructions from text into pictures with NLP and computer graphics techniques. The BabelDr system (Bouillon et al., 2021) proposes pictographs and allows to translate spoken medical utterances in various languages to communicate with migrants and deaf patients in hospitals. For people with ID, de Knegt et al. (2016a,b) designed a tool with pictographs, called STOP-ID, to aid the self-reporting of pain (affect, location, intensity and quality). The authors also tested the ability to recognize representations for vocabulary and pain of their tool in adults with ID.

The comprehension of single pictographs in context is increasingly evaluated with users: e.g. for the patient responses to medical questions (Norré et al., 2021a). However, most studies still rely on automated metrics used in MT such as BLEU

[1] https://arasaac.org

Ninth Workshop on Speech and Language Processing for Assistive Technologies (SLPAT-2022), pages 44 - 49

(Papineni et al., 2002), NIST (Doddington, 2002), etc. to assess sentences automatically translated into pictographs (Sevens, 2018; Vaschalde et al., 2018; Norré et al., 2021b). Mihalcea and Leong (2008) tested sentences whose nouns and verbs had been automatically translated into pictures. Finally, some evaluations are also carried out by researchers, such as in Bui et al. (2012), which rated the correctness of 49 patient instructions converted with Glyph. More recently, Bulté et al. (2021) manually evaluated the comprehension and the lexical coverage of sentences generated into three pictographic languages by their translation system.

3 Translation System

Our system was originally designed for the online communication of people with ID and was hence optimised for social media context (Sevens, 2018). It translates texts written in four natural languages into any combination of four pictograph sets. This paper focuses on French and the medical domain with Arasaac pictographs (see Section 4), as there are fewer medical pictographs in the other sets.

The text to translate first undergoes shallow linguistic analysis: sentence detection, tokenization, POS-tagging and lemmatization with TreeTagger (Schmid, 1994), simple detection of multi-word expressions (MWE), processing of specific French phenomena based on rules and dictionaries (Norré et al., 2021b). Then, each word of the text can be translated through two routes: the semantic route and the direct route. In the semantic route, each word is looked up in the WOLF database (Sagot and Fišer, 2008), a French version of WordNet (Miller, 1995). If it is not found, hyperonym and antonym relations of WOLF are used to get substitute translation. For example, as there is no pictograph for *saumon* (salmon), the word is translated by its hyperonym *poisson* (fish). The word *infecter* (infect) does not have a pictograph and is translated by its antonym followed by the negative pictograph, *désinfecter non* (desinfect no). For the direct route, we build a dictionary for the pictographic language that contains the words not covered by WOLF (e.g., prepositions, pronouns, etc.). Pictograph filenames (i.e. French lemmas) are linked to their identifiers available on the Arasaac website. To choose the optimal path while converting a sequence of lemmas to a sequence of pictographs, we use a search algorithm A* (Vandeghinste et al., 2015).

Compared to our previous work (Norré et al., 2021b), various improvements were brought to our system. We updated our pictograph database with new pictographs from Arasaac API,[2] as more medical pictographs have been added due to the Covid pandemic. Several AAC systems with pictographs use a color coding system that informs about the syntactic category of the words represented. This makes it possible to improve the learning of vocabulary and therefore its use. We implemented the coding system of Fitzgerald (1949), which highlights the borders of pictographs with colors, depending on their POS: green for verbs, blue for adjectives, etc. At the beginning of sentences, we also generate a temporal pictograph for past and future tenses (see Figure 1), as Sevens et al. (2017). We added a WOLF relation: eng_derivative, to get similar concepts with a different POS tag, e.g. for the adjective *respiratoire* (breathing), our system translates it by the verb *respirer* (breathe). It is the equivalent of xpos_near_synonym relation in other WordNets. We will therefore call it the xpos relation in Section 5. Finally, we added simple rules of compression for different French phenomena found in our previous evaluation (deletion of some function words, auxiliaries, verb-subject inversion in questions, simplification of some imperative structures for patient instructions). These rules are based on an analysis of our system's output, the advice from a speech therapist and a syntactic analysis carried out on medical sentences from the BabelDr system with the Berkeley Neural Parser (Kitaev and Klein, 2018).[3]

4 Translation Strategies for Pictographs

As we aim to automatically translate physician questions and patient instructions into Arasaac pictographs, we ran into the issue of the lack of a large authentic medical corpora in pictographs built by AAC users and the lack of translation guidelines to create such as a corpus. Therefore, we have investigated the actual strategies used by speech therapists to carry out such translation. This section first describes the data to translate, then lists the different translation strategies observed.

4.1 Data Set

For our translation experiment, we used the Arasaac pictograph set, an open source set that is in-

[2] https://arasaac.org/developers/api
[3] https://github.com/nikitakit/self-attentive-parser

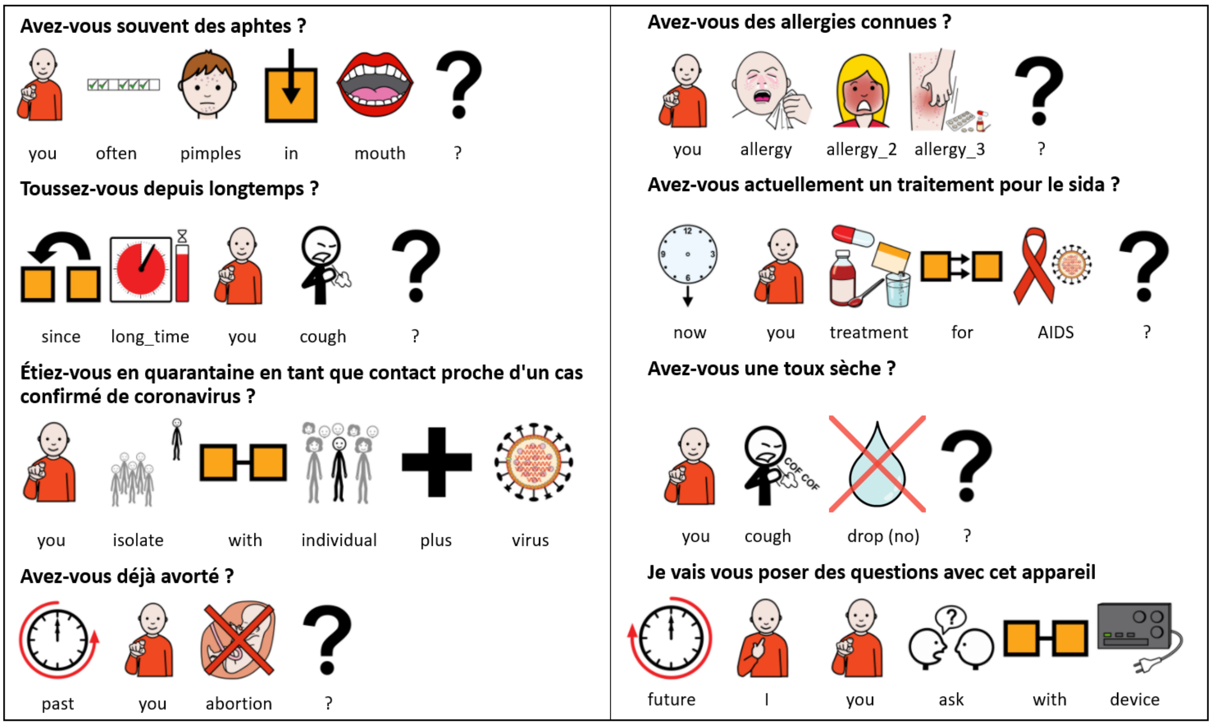

Figure 1: Examples of possible medical translations into Arasaac pictographs.

creasingly used by disabled people. This database includes over 12,000 pictographs in colours (also available in black and white and customizable online). Many domains (Paolieri and Marful, 2018), including communication in health sciences, are represented in this set. In November 2021, there were 1,126 medical pictographs grouped into 45 (sub)categories, such as medical procedures, covid-19, symptoms, etc. For the medical data to translate, we used French sentences of the BabelDr translation system (Bouillon et al., 2021), designed to facilitate communication between physicians and allophone patients. These data include physician questions, patient instructions, and greetings.

4.2 Translation Strategies

As Vandeghinste and Schuurman (2014) noted, a pictograph translation is not a literal translation. Various translation strategies can be used, the simpler one consisting in looking up each word lemma to translate in the pictograph set (Vaschalde, 2018). To uncover more sophisticated strategies, we asked a retired speech therapist – trainer and director of a Belgian AAC association – to manually translate 100 sentences from our medical data into Arasaac pictographs.

By comparing the original text and these sentences translated into pictographs, we noted at least 10 operations to improve the translation and the lexical coverage if a pictograph (filename) is missing: 1) deletion: delete some words of source sentence (e.g., articles, auxiliaries, etc.);[4] 2) insertion: if there is no pictograph for a (technical) term, insert a paraphrase with general concepts more easily comprehensible for patient (*aphtes: boutons dans bouche* | canker sores: pimples in mouth, as in Strasly et al. (2018) for translating sign language) or insert a clarification (*allergies connues: allergie allergie_2 allergie_3* | known allergies: allergy allergy_2 allergy_3); 3) moving: move one or several words (*toussez-vous depuis longtemps ?: depuis longtemps vous toussez ?* | cough you since long time ?: since long_time you cough ?); 4) synonym: replace by a synonym with an identical POS (*actuellement: maintenant* | currently: now); 5) hyperonym (*coronavirus: virus*); 6) hyponym (*personne: individu* | person: individual); 7) antonym (*sèche: goutte [non]* | dry: drop [no]); 8) POS change (*avorter: avortement* | abort: abortion); 9) compound2single: replace a MWE by a single word (*poser des questions: demander* | ask questions: ask); 10) replacement: replace one word with another with different root and/or POS, not by a synonym/hyperonym/etc. (*quarantaine: isoler* |

[4] The deletion depends on skill of user with ID. Note that removing all words by POS (e.g., adverbs) can change the meaning of sentences a lot (Vaschalde et al., 2018).

quarantine: isolate).

Figure 1 shows examples of expected medical translations into Arasaac pictographs.[5] Several translation operations can be combined in a sentence. These operations are already partially taken into account in the French version of the system we described in the previous section.

5 Preliminary Evaluation

We present the system tuning and an automated evaluation (Section 5.1), before the manual evaluation to assess the medical coverage (Section 5.2).

5.1 System Tuning and Automated Evaluation

For tuning and evaluation purposes, 150 additional medical sentences were manually translated into Arasaac by the authors. 60 sentences were used to tune the hyperparameters of the system (Vandeghinste et al., 2015) – related to WOLF relations, pictograph features and route preference – with a local hill climbing algorithm (5 trials of 50 iterations) using the BLEU metric (Papineni et al., 2002) as Norré et al. (2021b) on an email corpus.

	BLEU	WER	PER
- xpos relation	30.3 (2.3)	55.5 (2.1)	50.5 (1.9)
+ xpos relation	27.3 (2.2)	61.4 (2.6)	56.3 (2.5)

Table 1: System results on medical data for Arasaac: BLEU, WER and PER metrics (mean and std. dev.).

Then, we automatically evaluated the translation system on the remaining 90 sentences (Table 1). The BLEU scores are in line with our study (Norré et al., 2021b). For the French Text-to-Picto system, we got a BLEU score of 31.3 on a medical corpus for Arasaac, but with a largest reference corpus in which all the words had to be translated, including function words. Adding the xpos relation in our system (see Section 3) does not improve the results.

5.2 Manual Evaluation

Two experiments were carried out. We first calculated the number of untranslated words on 700 sentence transcripts of real physician questions, recorded with speech recognition of BabelDr system (Bouillon et al., 2021). We also evaluated 700 sentences, called canonicals, linked to each of these transcripts in the BabelDr system. Table 2 shows the number of untranslated types (and untranslated tokens in brackets). The use of xpos relation allows to translate more words even if we did not evaluate if all these translated words were correct.

	Transcripts	Canonicals
- xpos relation	126 (191)	102 (229)
+ xpos relation	110 (159)	81 (203)

Table 2: System results on medical data for Arasaac: number of untranslated types (and untranslated tokens).

As regards the lexical coverage, two authors of this paper manually evaluated 50 canonicals automatically generated with the system (without the xpos relation). They used UMLS concepts (Bodenreider, 2004) linked to these sentences to judge if the meaning was preserved. For each of the 103 concepts,[6] they annotated if the concepts were correctly translated into pictographs and by what type of representation (synonym, hyperonym or generic, hyponym or specific and polyseme). The Cohen's κ (Cohen, 1960) is 0.65, indicating that the agreement between both raters is substantial.

	Annotator 1	Annotator 2
Correct translation	62.1 (75.4)	71.8 (82.8)
No translation	37.8 (24.5)	28.1 (17.1)

Table 3: System results on medical data for Arasaac: lexical coverage (in %).

Table 3 shows the results of medical coverage by UMLS concept. The most used relation is synonymy. The annotators reported 5-6 hyperonyms (*nausée|diarrhée: symptomes* | nausea|diarrhea: symptoms), 2-5 hyponyms (*examen: examen des yeux* | examination: eye examination) and 3-4 polysemous words (*enceinte* means pregnant or speaker in French). There were also some MWE (*prise de sang* | blood test) incorrectly translated by two pictographs (*tenir sang* | grasp blood), but the MWE we had annotated with two synsets were correctly translated by a single pictograph (*mal à la tête* | headache). The untranslated words were mainly adjectives (*régulier* | regular) – more difficult to represent – and nouns (*type* | type). Figure 2 shows examples of system's outputs in Arasaac pictographs. The medical coverage can be still improved, especially the precision of the system, e.g. testing other lexical resources or NLP techniques that exploit the translation strategies into pictographs.

[5]The glosses are given using the English filenames of Arasaac pictographs. We added underscores and numbers.

[6]Or 187 concepts if we include the no UMLS concepts (e.g., pronouns and question marks). We also give results on this total in brackets in the table.

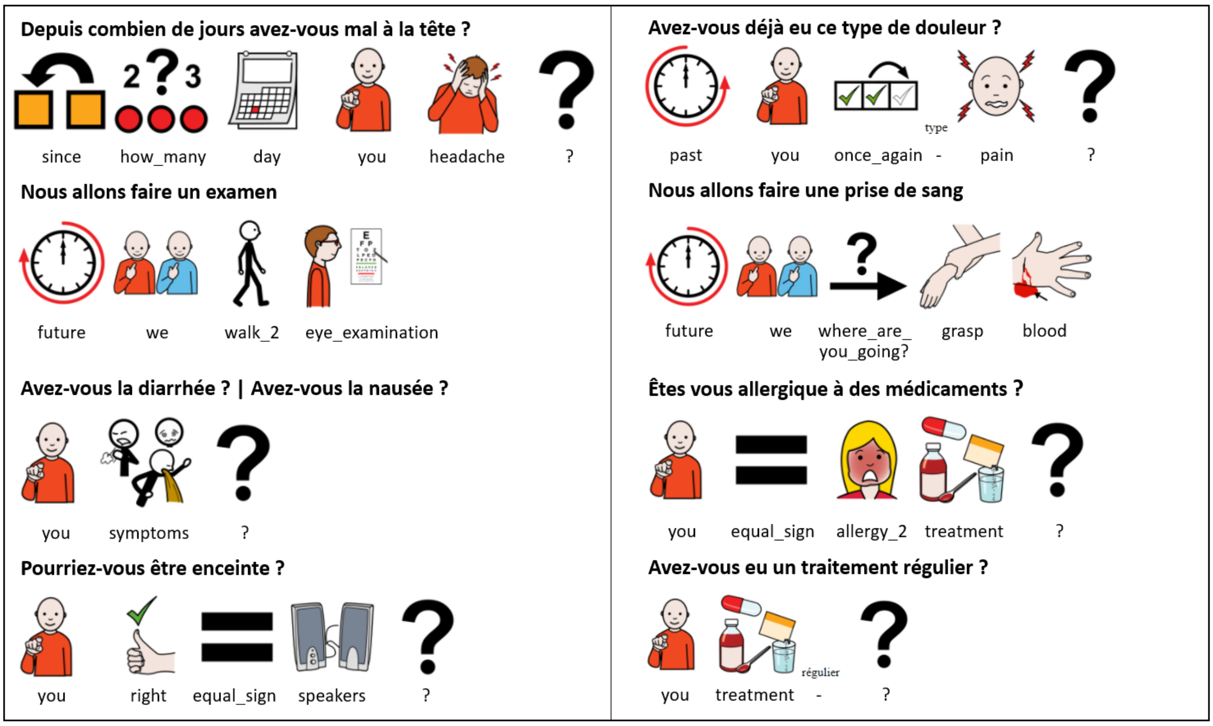

Figure 2: Examples of system's outputs in Arasaac.

6 Conclusion

We proposed an original way to investigate the medical coverage of our translation system from French into Arasaac pictographs using UMLS concepts. We also discussed the translation strategies into pictographs for medical sentences.[7] There is room for further improvement to specialize this system to the medical dialogue between physician and patients with ID. Some linguistic phenomena are not yet taken into account in the French system, such as the word sense disambiguation. Both the pictographic representations and the sentence comprehensibility into pictographs by the target users would need to be further investigated in context.

Acknowledgements

This work is also part of the PROPICTO project, funded by the Fonds National Suisse (N°197864) and the Agence Nationale de la Recherche (ANR-20-CE93-0005). The pictographs used are property of the Aragon Government and have been created by Sergio Palao to Arasaac. Aragon Government distributes them under Creative Commons License.

[7]The source code of the Text-to-Picto system, the manual and automated translations in Arasaac pictographs are available for the research community at the following address: https://github.com/VincentCCL/Picto.

References

Juliana Alvarez. 2014. Visual design. A step towards multicultural health care. *Arch Argent Pediatr*, 112(1):33–40.

David R. Beukelman and Pat Mirenda. 1998. *Augmentative and alternative communication*. Paul H. Brookes Baltimore.

Olivier Bodenreider. 2004. The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic acids research*, 32(suppl_1):D267–D270.

Pierrette Bouillon, Johanna Gerlach, Jonathan David Mutal, Nikolaos Tsourakis, and Hervé Spechbach. 2021. A speech-enabled fixed-phrase translator for healthcare accessibility. In *Proceedings of the 1st Workshop on NLP for Positive Impact*. Association for Computational Linguistics.

Duy Duc An Bui, Carlos Nakamura, Bruce E. Bray, and Qing Zeng-Treitler. 2012. Automated illustration of patients instructions. In *AMIA Annual Symposium Proceedings*, volume 2012, page 1158. American Medical Informatics Association.

Bram Bulté, Vincent Vandeghinste, Leen Sevens, Ineke Schuurman, and Frank Van Eynde. 2021. Can pictograph translation technologies facilitate communication and integration in migration settings? *Computational Linguistics in the Netherlands Journal*, 11:189–212.

Jacob Cohen. 1960. A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46.

Nanda de Knegt, Frank Lobbezoo, Carlo Schuengel, Heleen M. Evenhuis, and Erik J. A. Scherder. 2016a. Self-Reporting Tool On Pain in People with Intellectual Disabilities (STOP-ID!): a Usability Study. *Augmentative and Alternative Communication*, 32(1).

Nanda de Knegt, Carlo Schuengel, Frank Lobbezoo, Corine M. Visscher, Heleen M. Evenhuis, Judith A. Boel, and Erik J. A. Scherder. 2016b. Comprehension of pictograms for pain quality and pain affect in adults with Down syndrome. *Journal of Intellectual & Developmental Disability*, 41(3):222–232.

George Doddington. 2002. Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In *Proceedings of the second international conference on Human Language Technology Research*, pages 138–145.

Edith Fitzgerald. 1949. *Straight language for the deaf: a system of instruction for deaf children*. Volta Bureau.

Antony Akash Janakiram, Pierrette Bouillon, Johanna Gerlach, Patricia Martha Hudelson Perneger, and Hervé Spechbach. 2021. J'ai de la peine à communiquer avec mon patient aux urgences. Quels sont les outils disponibles ? *Revue médicale suisse*, 7(739):995–998.

Nikita Kitaev and Dan Klein. 2018. Constituency parsing with a self-attentive encoder. *arXiv preprint arXiv:1805.01052*.

Rada Mihalcea and Chee Wee Leong. 2008. Toward communicating simple sentences using pictorial representations. *Machine translation*, 22(3):153–173.

George A. Miller. 1995. WordNet: A lexical database for English. *Communications of the ACM*, 38(11).

Ankita Nandy. 2019. Beyond words: Pictograms for Indian languages. *International Journal of Research in Science and Technology*, 9(1):19–25.

Magali Norré, Pierrette Bouillon, Johanna Gerlach, and Hervé Spechbach. 2021a. Evaluating the comprehension of Arasaac and Sclera pictographs for the BabelDr patient response interface. In *Proceedings of the 3rd Swiss Conference on Barrier-free Communication*, pages 55–63. ZHAW Zurich University of Applied Sciences.

Magali Norré, Vincent Vandeghinste, Pierrette Bouillon, and Thomas François. 2021b. Extending a Text-to-Pictograph System to French and to Arasaac. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2021)*, pages 1050–1059.

Daniela Paolieri and Alejandra Marful. 2018. Norms for a Pictographic System: the Aragonese Portal of Augmentative/Alternative Communication (ARASAAC) System. *Frontiers in psychology*, 9:2538.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. BLEU: A method for automatic evaluation of machine translation. In *Proceedings of the 40th annual meeting of the Association for Computational Linguistics*, pages 311–318.

Benoît Sagot and Darja Fišer. 2008. Building a free French WordNet from multilingual resources. In *OntoLex*, Marrakech, Morocco.

Helmut Schmid. 1994. Probabilistic part-of-speech tagging using decision trees. In *New methods in language processing*, page 154.

Leen Sevens. 2018. *Words Divide, Pictographs Unite: Pictograph Communication Technologies for People with an Intellectual Disability*. LOT, JK Utrecht, The Netherlands.

Leen Sevens, Vincent Vandeghinste, Ineke Schuurman, and Frank Van Eynde. 2017. Simplified text-to-pictograph translation for people with intellectual disabilities. In *International Conference on Applications of Natural Language to Information Systems*, pages 185–196. Springer.

Irene Strasly, Tanya Sebaï, Evelyne Rigot, Valentin Marti, Jesus Manuel Gonzalez, Johanna Gerlach, Hervé Spechbach, and Pierrette Bouillon. 2018. Le projet babelDr : rendre les informations médicales accessibles en Langue des Signes de Suisse Romande (LSF-SR). In *Proceedings of the 2nd Swiss Conference on Barrier-free Communication*, pages 92–96.

Vincent Vandeghinste and Ineke Schuurman. 2014. Linking pictographs to synsets: Sclera2Cornetto. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, volume 9, pages 3404–3410. ELRA, Paris.

Vincent Vandeghinste, Ineke Schuurman, Leen Sevens, and Frank Van Eynde. 2015. Translating text into pictographs. *Natural Language Engineering*, 23(2):217–244.

Céline Vaschalde. 2018. Génération automatique de pictogrammes à partir de la parole pour faciliter la mise en place d'une communication médiée. Master's thesis, Université d'Orléans.

Céline Vaschalde, Pauline Trial, Emmanuelle Esperança-Rodier, Didier Schwab, and Benjamin Lecouteux. 2018. Automatic pictogram generation from speech to help the implementation of a mediated communication. In *Proceedings of the 2nd Swiss Conference on Barrier-free Communication*.

Krzysztof Wołk, Agnieszka Wołk, and Wojciech Glinkowski. 2017. A cross-lingual mobile medical communication system prototype for foreigners and subjects with speech, hearing, and mental disabilities based on pictograms. *Computational and mathematical methods in medicine*, 2017.