Computational Linguistics in the Netherlands Journal 11 (2021) 189-212 Submitted 10/2021; Published 12/2021

Can Pictograph Translation Technologies Facilitate Communication and Integration in Migration Settings?

Bram Bulté* BRAM.BULTE@CCL.KULEUVEN.BE
Vincent Vandeghinste** VINCENT.VANDEGHINSTE@IVDNT.ORG
Leen Sevens* LEEN.SEVENS@KULEUVEN.BE
Ineke Schuurman* INEKE.SCHUURMAN@CCL.KULEUVEN.BE
Frank Van Eynde* FRANK.VANEYNDE@CCL.KULEUVEN.BE

* *Centre for Computational Linguistics, KU Leuven, Belgium*

** *Instituut voor de Nederlandse Taal, Leiden, The Netherlands*

Abstract

In this pilot study, we investigate the potential of pictograph translation technologies for facilitating communication and integration in the context of migration. We incorporate a new pictograph set in an existing text-to-pictograph translation system and carry out evaluations on three sets of authentic data (language classes, news articles, websites of local governments). We also evaluate whether a component targeting named entities can increase the coverage of the system. Our results show that, even though the pictograph translations can successfully represent parts of input sentences, conveying their full meaning proves to be a difficult task. We conclude that using the text-to-pictograph translation system as such in a migration context is not recommended. At the same time, we suggest other potential applications for the system in a migration context and point to potential improvements.

1. Introduction

The number of international migrants worldwide is currently estimated at around 272 million, or approximately 3.5% of the world's population (IOM 2019). In 2019 alone, 2.7 million people immigrated to the European Union, and 699,000 persons from nearly 150 countries applied for asylum in the EU.[1] In the same year, almost 300,000 asylum seekers received a form of protection status in one of the EU member states. In Belgium, around 20,000 persons apply for asylum each year (Fedasil 2019). Their backgrounds are very diverse, with people originating from countries such as Afghanistan, Palestine, Guinea, Eritrea, Somalia and Syria. They also speak a wide range of first and second languages, amongst which Arabic, French, English, Pashto, Farsi and Spanish are, currently, the most frequent ones. In addition, there is considerable variation when it comes to literacy and educational background, with around 10% of asylum seekers being illiterate, and an additional 40% being functionally illiterate (Hooft et al. 2020).

It should come as no surprise that there are considerable communication challenges in a linguistically highly diverse migration context, both with respect to official communication between asylum seekers and government (officials), and in the context of more informal day-to-day exchanges and activities, such as grocery shopping, public transport use, newspaper reading, dealing with school administration or maintaining social relationships (Hooft et al. 2020). Such challenges are not easy to overcome, as, for example, producing and providing tailor-made communication is time-consuming, interpreting services are not always available, learning a second language can be a lengthy process, and available technological tools have clear limitations. Fedasil, the Belgian Federal Agency for the Reception of Asylum Seekers, currently provides written information in 14 languages on its

1. Eurostat asylum statistics, `https://ec.europa.eu/eurostat/statistics\-explained/index.php?title=Asylum_statistics`

information hub,[2] which also includes text-to-speech options. Information within asylum centres is also displayed and provided in a wide range of languages. Written language is not accessible to everyone, however, and it is practically impossible to cover all necessary languages. Tools such as machine translation (e.g., Google Translate) in combination with text-to-speech are sometimes used by migrants, but these are also not accessible to everyone, for example due to differing levels of (digital) literacy, and the produced translations can lead to confusion and misunderstandings (Hooft et al. 2020).

In this pilot study, we combine two of the tools that are already occasionally used to help overcome communication challenges, namely pictographs and machine translation. Seeing that pictograph translation is used successfully as a communication tool for people with intellectual disabilities in Flanders (Sevens et al. 2018), *De Rand*,[3] an organisation funded by the Flemish government to promote, amongst other things, the integration of non-Dutch-speaking persons in the communes around Brussels, wanted to know whether the same technology could be useful in a migration context. This question seemed justified, especially considering that they already make available printed booklets with different categories of pictographs to migrants, which are not always very user-friendly. We therefore set out to investigate the potential of pictograph translation technologies for facilitating communication and integration in migration settings, focusing on the Flemish context. We incorporate the *De Rand* pictograph set in an existing text-to-pictograph translation system and carry out evaluations on three sets of authentic data. We also evaluate whether a component targeting named entities can increase the coverage of the system, since these are not covered by the pictograph sets. As our initial tests showed that the generated pictograph translations did not sufficiently cover the meaning of the original sentences, we did not consider it opportune to involve the intended end-users in our evaluations, as was originally planned. The evaluations reported in this paper are, therefore, admittedly rather limited, but also, as we will argue, appropriate and valid.

This paper is structured as follows: section 2 presents the background to the study. This is followed by a description of the pictograph translation system (section 3) and the methodology and results of the experiments that were carried out (section 4). The findings of the study are discussed in section 5, and conclusions are drawn up in section 6.

2. Research background

We first discuss communication-related challenges and needs of migrants in Flanders (2.1), before turning to the role of pictographs in facilitating communication (2.2). In the final subsection, we provide a brief overview of existing tools for automated text-to-pictograph translation (2.3).

2.1 Communication-related challenges and needs of migrants in Flanders

In Flanders and elsewhere in Europe, language barriers for asylum seekers and refugees have been attested with regard to, amongst other things, healthcare (Cox 2017, Jensen et al. 2013, Krystallidou et al. 2021, Manirankunda et al. 2012, Pavli and Maltezou 2017), education (Koehler and Schneider 2019, Stevenson and Baker 2018), access to housing (Hamann and El-Kayed 2018), employment (Auer 2018, Lochmann et al. 2019) and entrepreneurship (Wauters and Lambrecht 2008). In a detailed case study of Belgian asylum centers, Hooft et al. (2020) show how such issues are compounded by varying levels of literacy. On the basis of interviews with asylum seekers and caregivers, they provide an overview of very basic and more fundamental problems that asylum seekers with low literacy face, as well as strategies and tools to potentially overcome these problems.

While their application is being evaluated, most asylum seekers in Belgium stay in one of around 60 asylum centres (Fedasil 2019). Asylum centres organise interpreting services to facilitate the communication between asylum seekers and social workers, legal experts and caregivers, and ad hoc

2. https://www.fedasilinfo.be
3. https://www.derand.be/

interpreting (by volunteers, often also other asylum seekers) is sometimes available. These options are, however, not helpful or accessible for everyone, nor at all times. For this reason, asylum seekers often rely on technological tools such as GPS, voice messages, digital pictures, predictive typing and Google translate (with text-to-speech). Even though translation tools are frequently used, caregivers noted that they often lead to misunderstandings (Hooft et al. 2020).

With the aim of improving communication and facilitating the integration of asylum seekers in society, Dutch or French language classes are organised in the asylum centres, mostly by volunteers. About one third of asylum seekers in Belgium are younger than 18, and for them education is compulsory. These children and adolescents often spend one or two years in a programme dedicated to learning Dutch (in so-called "OKAN" classes in Flanders and Brussels) or French ("DASPA" in Wallonia and Brussels), before being integrated in mainstream education. The diverse language backgrounds and literacy levels pose specific challenges for the organisation of these language classes.

2.2 Pictographs as a tool for facilitating communication

Pictographs can, perhaps somewhat naively, be conceived of as a sort of universal language, conveying messages without relying on speech, gestures or writing. Many potential advantages of pictograph communication have been attested. Pictographs are easy to learn (Tanimoto 1997) and pose a reduced cognitive load on their users, especially when compared to writing (McCoy 1998). This is, in part at least, due to the fact that they are less affected by the irregularities and inconsistencies that characterise natural languages (Leemans 2001). Moreover, to put it in Saussurean terms, with pictographs there is often a motivated link between the signifier and what it signifies (in contrast to the largely arbitrary nature of this connection in natural languages). They also make it possible to convey a great deal of information using limited space (Böcker 1996).

There are, however, a number of important caveats. First, pictograph sets are often limited in size and scope, as they mostly constitute fixed sets, developed in a particular context and with a specific purpose in mind (Sevens 2018). They can therefore not be used to express all concepts, entities and actions. Second, pictographs are best suited for representing concrete objects and actions. It is doubtful whether they can easily express abstract concepts and complex, potentially compounded semantics in a straightforward and unambiguous way. Third, with the exception of some very restricted domains, there is no established "standard" universal pictograph language. Fourth, even though pictographs may be easy to learn, efficient pictograph communication still requires training (Martínez-Santiago et al. 2015, Romski et al. 2009). Finally, each individual may interpret pictographs differently, for example due to differences in culture, ethnicity or social background (Goonetilleke et al. 2001, Huer 2000, Munemori et al. 2010). It can, therefore, not be taken for granted that pictographs are understood or correctly interpreted by users (Cowgill and Bolek 2003). It is clear that, especially in a highly diverse context, misunderstandings are to be expected (Soares 2015).

In spite of these potential shortcomings, pictographs are used for facilitating or augmenting communication in a wide variety of restricted domains and applications, such as warning or safety labels (Caffaro and Cavallo 2015, Davies et al. 1998), pharmacy (Dowse and Ehlers 1998, Montagne 2013), healthcare (Barros et al. 2014), (international) travel and transportation, traffic, personal computers and software / human-computer interaction (icons), and recently also in electronic communication in general (e.g., emoticons and emojis). They have also been proven useful in the context of augmentative and alternative communication, for example for enabling or facilitating communication with people with intellectual disabilities (Keskinen et al. 2012, Paolieri and Marful 2018).

In the migration context under investigation here, pictographs are used throughout asylum centres to facilitate communication (e.g., during courses, to establish 'house rules' or for announcing activities), and their use is recommended by the government for local administrations[4] and for

4. https://www.vlaanderen.be/publicaties/van-opvang-naar-samenleven

caregivers or social workers who experience barriers in communication.[5] Fedasil also extensively uses pictographs to facilitate navigation on its website.[6] *De Rand* freely provides booklets with pictographs grouped according to different categories such as food, shopping, sports and transportation, and also makes these pictographs available on their website.[7] Visual materials, including pictographs, are also frequently used in (beginner) second language classrooms and course books, especially when the learners come from diverse linguistic backgrounds and have little or no experience with reading and writing. Primary and secondary schools also sometimes augment texts with pictographs when communicating with the parents of pupils who do not speak the language that is used in the school.

2.3 Automated text-to-pictograph translation

Whereas the use of pictographs is rather widespread, only a handful of attempts have been made to design a system that automatically converts text into pictographs, and vice versa (Sevens 2018). Options to augment text with automatically retrieved images have been explored more extensively (Joshi et al. 2006, Poots and Bagheri 2018). We will first look at some examples of tools that enrich text with pictures, before turning to systems that convert text into pictographs. In our overview we do not include systems that generate new images from text, such as WordsEye (Coyne and Sproat 2001), as this would go beyond the scope of the present article. Instead, we focus on those systems that make use of a database of existing images.

Joshi et al. (2006) designed a tool for automated story picturing. After keywords are extracted from the story, related illustrations are retrieved from a database of annotated pictures. WordNet (Miller 1995) is used to assess the lexical similarity of the keywords, and a predefined number of pictures is selected on the basis of reinforcement-based ranking. Agrawal et al. (2011) developed a set of algorithms to automatically retrieve images to illustrate textbooks. To this end, they made use of an internet search engine, restricted to Wikipedia pages. On the basis of a user study they concluded that the system was able to propose relevant images that aided comprehension of the textbook contents. The system described by Mihalcea and Leong (2008) takes text as input, and generates a combination of textual and pictorial output. The images are retrieved from a database containing both photographs and illustrations, which are linked to certain (basic) nouns and verbs. Their experiments confirmed that the converted messages could be understood reasonably well.

Turning to systems that work exclusively with pictographs, Müller et al. (2010) tested a text messaging tool for people with communicative disabilities. One of its features is that it can convert received text messages into pictographs, by making use of a database in which individual words are coupled with their pictorial counterparts. They showed that using the tool increased their participants' ability to communicate remotely. Bautista et al. (2017) and Martín et al. (2018) describe a system to translate Spanish text into pictographs using the ARASAAC[8] database, which was created specifically for augmentative and alternative communication. After linguistic analysis, the database is queried for pictographs representing both single words and multi-word expressions, from which a pictograph translation is generated. There are also a number of commercial tools available for generating pictographs on the basis of text, such as Widgit Symbols.[9]

In this study, we work with the text-to-pictograph translation system developed by Vandeghinste et al. (2017) and Sevens (2018). We use the Dutch version of the tool, which is otherwise also available in Spanish and English (Sevens et al. 2015). It has recently also been adapted to cover French text input, with ARASAAC pictographs as output (Norré et al. 2021). It is important to note that this system was originally developed to facilitate the (online) communication of people with intellectual disabilities and was hence optimised for this specific context and domain. Moreover, its original

5. https://www.integratie-inburgering.be/nl/communicatiewaaier
6. https://www.fedasilinfo.be
7. https://www.pictogrammendatabank.be
8. https://www.arasaac.org
9. https://www.widgit.com

end-users were trained to communicate with the pictographs that are implemented in the system. They also shared the same language background. As opposed to that, in this study we focus on the use of the translation tool in a migration setting. Here, the intended users come from highly diverse cultural and linguistic backgrounds, and have varying levels of literacy. Our main aim is to investigate whether the tool can be useful when it is used directly by migrants themselves to help understand Dutch text. There are, of course, other potential use cases as well, which involve, for example, communication between caregivers and migrants,[10] teachers creating materials for language classes, or government bodies wishing to provide information that is more accessible.

3. System description

A detailed description of the original text-to-pictograph translation system is provided by Vandeghinste et al. (2017) and Sevens (2018). The source code is available on Github[11] and a public demo is also available online.[12] In this section, we provide a short overview of the different system components. First, we describe the three pictograph sets the system works with (3.1) and the use of WordNets (3.2). The following subsection presents the component that was added to deal with named entities (3.3). Finally, the text-to-pictograph pipeline is described in the last subsection (3.4).

3.1 Pictograph sets

3.1.1 Sclera and Beta

For the original translation system, two pictograph sets were used: *Sclera*[13] and *Beta*.[14] Both *Sclera* and *Beta* pictographs were specifically developed to facilitate communication with people with intellectual disabilities.

There are over 13,000 *Sclera* pictographs, most of which are in black and white. The majority of these pictographs cover nouns, verbs and adjectives. More than half of the *Sclera* pictographs represent complex meanings corresponding to several Dutch words (e.g., verb+object). Examples of *Sclera* pictographs are shown in Figure 1.

Figure 1: *Sclera* pictographs for *trein* ('train'), *warm* ('warm') and *kat eten geven* ('feed the cat')

The roughly 3,500 *Beta* pictographs are in colour, and mostly represent single lexical units. Compared to *Sclera*, they cover a wider variety of adjectives, adverbs and prepositions. Figure 2 provides examples of *Beta* pictographs.

3.1.2 De Rand

In the current version of the translation system we also added the *Rand* pictograph set. This set of pictographs was designed for communication in migration contexts in Flanders, with the

10. See Norré et al. (2021) for an example of pictograph translation use in a medical context.
11. https://github.com/VincentCCL/Picto
12. http://picto.ccl.kuleuven.be
13. https://www.sclera.be
14. https://www.betasymbols.com

Figure 2: *Beta* pictographs for *boot* ('boat'), *koud* ('cold') and *voeren* ('feed')

aim of assisting people learning Dutch by providing visual support to facilitate the understanding of words/messages they are confronted with in everyday life. As mentioned in section 2.2, these pictographs are made available by the organisation *De Rand* in the form of booklets and through a web interface. The pictographs are in colour, and they mainly represent single words. Most pictographs in this set cover nouns, but there is a fair number of verbs as well. Adjectives are less frequent in this set. Whereas the *Sclera* and *Beta* pictographs have fixed dimensions, the aspect ratio of *Rand* pictographs is variable. Examples of *Rand* pictographs can be found in Figure 3.

Figure 3: *Rand* pictographs for *vliegtuig* ('plane'), *groot* ('large/big/tall') and *waaien* ('blow')

3.2 WordNet/Cornetto

3.2.1 Pictograph linking

The pictographs from the three sets were manually linked to synonym sets (*synsets*), as in Vandeghinste and Schuurman (2014), in Cornetto, a Dutch version of WordNet (Vossen et al. 2013). In total, 5,710 *Sclera*, 2,760 *Beta* and 2,164 *Rand* pictographs were linked to their corresponding synsets. Not all pictographs were retained since there are often different pictographs representing the same synonym set, and some pictographs were deemed to be too specific (e.g., restricted to the specific context in which they were developed). Because they are linked to groups of synonyms, the lexical coverage of the pictographs is increased compared to, for example, lemma-based matching (Mihalcea and Leong 2008, Müller et al. 2010).

3.2.2 Updates

In order to further increase the coverage of the system, 120 Flemish words were added to the Dutch Wordnet, which mainly contains standard Dutch vocabulary and regional variants of words that are used in the Netherlands. Examples of added words are *bompa* ('grandfather'), *plein* ('square') and *poetsvrouw* ('cleaning lady').[15]

This last example brings us to the second update that was implemented. WordNet/Cornetto synsets are often not detailed enough to properly deal with natural gender, as already discussed in Schuurman et al. (2020). For example, in both Cornetto and Open Dutch WordNet (Postma

15. As the system described here translates text into pictographs, and not the other way around, issues related to politically correct language use (i.e., which term to use?) are not discussed in this paper.

et al. 2016) there is one synset containing both *zanger* ('singer') and *zangeres* ('female singer'). Linking pictographs with this synset would mean that we are unable to control the pictograph that is generated. With this in mind, we created new synsets for *zanger* and *zangeres*, the first with as second hyperonym *man* ('man'), the other one with second hyperonym *vrouw* ('woman'). A similar example, involving the words *juf* ('female teacher') and *meester* ('male teacher'), is illustrated in Figures 4 and 5, which show the old and new structure of the related synsets, respectively.

Sometimes also in Dutch the same word is used for both sexes, while two pictographs are available. An example would be *bakker* ('baker'). In such a case we take this as a gender-neutral hyperonym, adding as hyponyms two new instantiations of *bakker*, once more adding these gender-specific hyperonyms. Note that the original 'gender-neutral' concepts are kept, in order to be used in neutral environments. In total, 380 lexical entries and 151 synonym sets were added in this way.

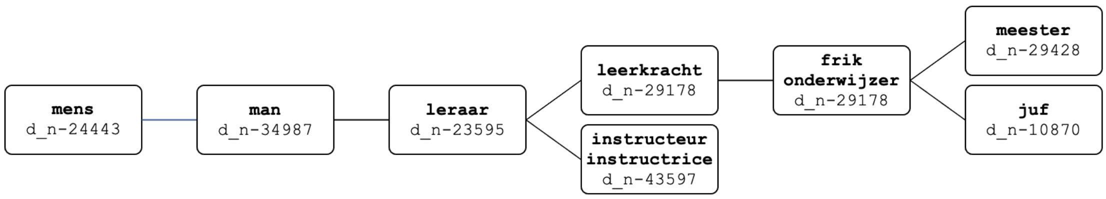

Figure 4: Synsets related to *juf* ('female teacher') and *meester* ('male teacher'), original hierarchy

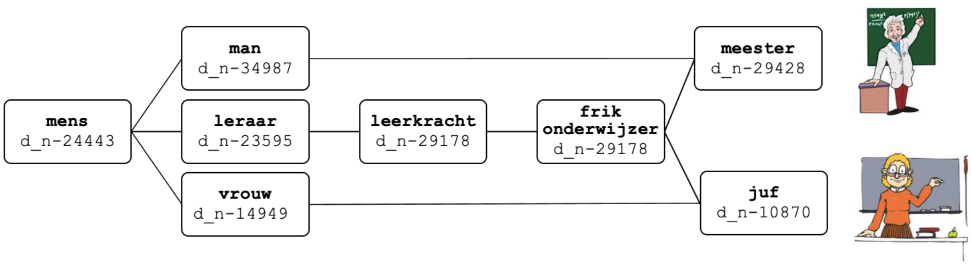

Figure 5: Synsets related to *juf* ('female teacher') and *meester* ('male teacher'), updated hierarchy with multiple inheritance + *Rand* pictographs

3.3 Named entities

The three pictograph sets described in section 3.1 provide little to no coverage of named entities (e.g., persons, places). We test a component that identifies named entities[16] and links them to pictures found on Wikipedia. To this end, we use a combination of Wikifier[17] (Brank et al. 2017) and the MediaWiki API.[18] Wikifier identifies potentially relevant Wikipedia pages for input text, following a pagerank-based approach. The pages are linked to one or several consecutive input tokens. The MediaWiki API allows us to retrieve the main or first image that appears in the selected Wikipedia article. The system developed by Agrawal et al. (2011) also used Wikipedia as a resource for retrieving images.

16. Note that we define *named entity* loosely at this point, as we do not perform named entity recognition as such, but rather attempt to find Wikipedia articles without restrictions on what these articles deal with.
17. http://wikifier.org
18. https://www.mediawiki.org

The text-to-pictograph translation system can be configured to only show images retrieved for tokens for which no pictographs were found, or to give priority to the images. In our experiments (see section 4) we evaluate all images that are retrieved. In addition, we carry out tests using two different thresholds for *pagerank score*,[19] a key parameter used in Wikifier's pagerank-based method. Setting this parameter higher results in more pages being retrieved (potentially favouring recall), whereas lowering it leads to a more restricted output (which should be better for precision). Based on preliminary testing as well as previous experiments in the field of lexical simplification (Bulté et al. 2018), we test two values for this parameter: 1 and 0.95.

3.4 Text-to-pictograph pipeline

3.4.1 Pre-processing textual input

The current pre-processing of the source language is an adaptation of the system described in Vandeghinste et al. (2017), now involving a deep linguistic analysis using the Alpino parser (van Noord 2006), instead of a shallow analysis. The system also contains a module for separable verb detection (Vandeghinste 2002), and allows for syntactic simplification (Sevens 2018).

The simplification module, among other things, disentangles embedded clauses and identifies interrogative and passive sentences. Syntactic constituents are reordered in order to produce active subject-verb-object (SVO) clauses.[20] Finally, as in Vandeghinste et al. (2017), words with negative polarity, such as *niet* ('not') and *geen* ('no'), are identified, and coupled with their heads.

3.4.2 Retrieving pictographs

A dictionary-style resource is used for a direct look-up of pictographs for a restricted set of words (e.g., pronouns). The remaining lexical items are linked to their corresponding Cornetto synsets, and all synsets for which a pictograph is available are retrieved. This linking is achieved by means of a lemma-based database look-up (Vandeghinste et al. 2017). In case multiple synsets are retrieved, word sense disambiguation, relying on support vector machines, is applied to attribute probabilities to the different synsets (Sevens et al. 2016).[21]

If a synset is retrieved that is not linked to a pictograph, the semantic relationships between Cornetto synsets are used to find semantically related pictographs. Three relations are used: hyperonymy (*HAS_HYPERONYM*), antonymy (*ANTONYM*), and near synonymy involving a different part-of-speech tag (*XPOS_NEAR_SYNONYM*). Multiple relations can be combined. An example of how hyperonymy is exploited would be the pictograph for the super-category *vis* ('fish') being retrieved for the word *aal* ('eel'), which has no specific pictograph attached to it. In case an antonym is found for a word without pictograph (e.g., *ill* for *recovered*), the pictograph for the antonym is retrieved together with the pictograph representing negation. Near synonyms with a different part-of-speech tag are, for example, nouns and adjectives (*female* and *feminine*), or nouns and verbs (*feeling* and *feel*).

Each semantic relationship is assigned a penalty. These penalties were tuned using a local hill climber. They are combined with the word sense disambiguation probabilities to find the optimal pictograph sequence (Sevens 2018, Vandeghinste et al. 2017).

3.4.3 Finding the optimal path

An A^* search algorithm finds the optimal path for all pictographs that are retrieved for each (potentially simplified) input sentence. This optimal path takes into account a number of penalties related

19. Parameter *pageRankSqThreshold*.
20. Considering the diverse language backgrounds of the users, other canonical word orders (e.g., SOV and VSO) are also possible, and the 'writing direction' of pictographs can be made reversible (left→right and right→left).
21. `https://www.github.com/cltl/svm_wsd`

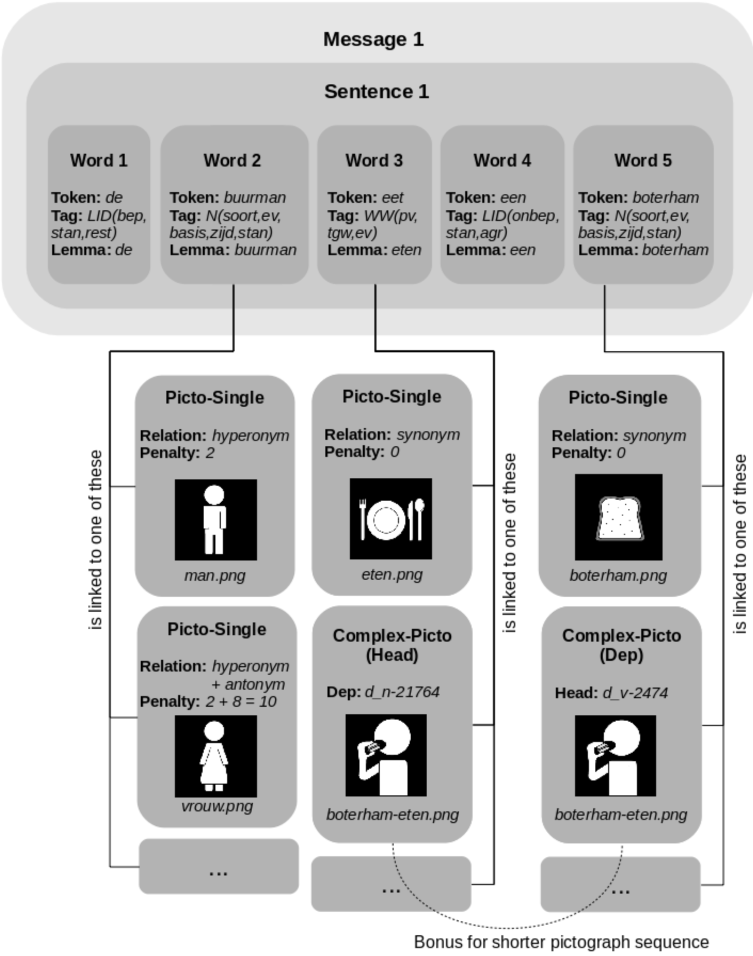

Figure 6: Translation of *De buurman eet een boterham* (‘The (male) neighbour eats a sandwich’) into *Sclera* pictographs (Sevens 2018, p. 62).

to the number and type of semantic relations between synsets involved in pictograph retrieval, as well as a weight for word sense disambiguation.

The text-to-pictograph translation process is illustrated in Figure 6 using the example sentence *De buurman eet een boterham* (‘The (male) neighbour eats a sandwich’). For *buurman*, pictographs are retrieved (with an associated penalty) for, amongst others, the hyperonym *man* and the hyperonym+antonym *vrouw*. For both *eet* and *boterham* a single ‘direct’ pictograph is retrieved. In addition, a pictograph is found that represents both words (*boterham-eten*). This ‘combined’ pictograph is more likely to be selected, since it represents a shorter pictograph sequence.

4. Experiments

We first provide the research questions in section 4.1. Section 4.2 details the methodology of the study, including data sets and evaluation procedures. The results of the experiments are provided in section 4.3.

4.1 Research questions

With the aim of fine-tuning the translation system and proposing helpful additions and/or changes, we evaluate its coverage and potential usability in a migration setting. We compare three authentic contexts (language classes, news, administration) using three pictograph sets (*Sclera*, *Beta*, *Rand*). The research questions can be formulated as follows:

- RQ1: Does the Dutch text-to-pictograph translation system produce comprehensible output at sentence level?
- RQ2: How many lexical items are covered by pictographs, and how useful are the retrieved pictographs?
- RQ3: How many named entities are covered by Wikipedia images, and how useful are these images?
- RQ4: How useful are additionally retrieved images, not linked to named entities?

Based on our literature review and previous experience with the text-to-pictograph translation system, we can formulate a number of hypotheses related to these research questions. We expect the task at hand to be (very) difficult for the translation system, given the relatively unrestricted nature of the chosen domains and the potential complexity of the messages that need to be translated. The pictograph sets are restricted in size, and certain concepts are difficult to express with pictographs in general, especially for users who do not have any relevant training. Moreover, the *Sclera* and *Beta* pictographs sets, as well as the original text-to-pictograph translation system, were designed for people with an intellectual disability, and not for the contexts investigated here. The *Rand* pictographs are aimed at beginner learners of Dutch, so they should be best suited for the language classes data set.

Generally speaking, we predict to find differences in performance between the three contexts, with beginner language classes probably representing the easiest use case, and administrative communication the most difficult. Factors such as sentence length and the degree of concreteness of lexical items will most likely have a considerable impact on the quality of the produced translations. These factors, however, will not be explicitly evaluated in the current study. We do not have specific predictions related to the retrieval of images for named entities.

4.2 Method

4.2.1 Data sets

The three contexts investigated in this pilot evaluation study, beginner Dutch language classes, newspaper articles and administrative communication, were identified as problematic by asylum seekers and their caregivers (Hooft et al. 2020). We collected 30 sentences for each context from randomly selected beginner-level course materials,[22] an easy-to-read online Dutch newspaper,[23] and websites of Flemish communes surrounding Brussels,[24] respectively. In all cases, sentences are taken from multiple documents. Whereas the texts in the first two data sets are clearly aimed at learners of Dutch, this is not the case for the administrative communication.[25]

Table 1 provides more information about the three data sets, focusing on the number of words, lexical units and named entities. The identification of lexical units is to some degree subjective, but

22. `https://www.taalunieversum.org/inhoud/nt2-beginnersdoelen/doos-op-rollen`
23. `https://www.wablieft.be`, also available as corpus (Vandeghinste et al. 2019).
24. `https://www.dilbeek.be`,`https://www.zaventem.be` and `https://www.vilvoorde.be`
25. There are government initiatives to promote the provision of accessible information and easy-to-read texts to citizens (see e.g., `https://overheid.vlaanderen.be/communicatie/heerlijk-helder`), but the corresponding guidelines are not (yet) always applied.

we tried to adhere to the definition of a lexical unit as "a unit of description made up of words and phrases" (Sinclair 1998, p. 23). Consequently, certain words are not counted as lexical units (e.g., articles, certain prepositions), while lexical units can consist of multiple words (e.g., *in staat zijn - to be able*). Similarly, defining and recognising named entities is not always a straightforward task (Nadeau and Sekine 2007, Marrero et al. 2013). For the purpose of our evaluation, we count proper nouns (referring to, for example, individuals, organisations or places) as named entities. We do not include temporal expressions or quantities.

	Language classes	**News**	**Administration**	**Total**
Sentences	30	30	30	90
Words	195	225	381	801
Average sentence length	6.5	7.5	12.7	8.9
Lexical units	153	164	244	561
Named entities	11	13	9	33

Table 1: Data set description

It is clear from this table that the administrative data contains the longest sentences, which are almost twice as long as those in the language classes data. The number of named entities is restricted in the three data sets, at least in part due to the nature of the texts, and we will therefore not split up our analyses related to image retrieval by data set. More details about the data sets, including sample sentences, are provided in Appendix A. The full data set is available online.[26]

4.2.2 Evaluation

We compare the three pictograph sets in terms of sentence comprehensibility (RQ1) and lexical coverage (RQ2). We also evaluate the component targeting named entities, and test two different similarity thresholds for picture retrieval (RQ3). Finally, we assess the potential usefulness of the retrieved images that are not linked to named entities (RQ4). It is important to note that named entities are not excluded from the calculation of the results for both sentence comprehensibility and lexical coverage, whereas they are not covered by the pictograph sets.

The evaluation in this pilot study is carried out by the researchers themselves, more specifically by the first author of this paper. The restrictive measures related to the Covid pandemic certainly played a role here, but preliminary results also led to our decision not to involve the target users at this stage of development. The evaluation procedure involves judgements that can be characterised as subjective, as concepts such as 'useful' and 'comprehensible' are clearly open to interpretation.

We use four categories to evaluate comprehensibility at the sentence level: 'no pictographs', 'not helpful', 'partially helpful' and 'potentially fully helpful'. 'No pictographs' simply means that no pictographs were retrieved for an input sentence. We use the classification 'not helpful' when we estimate that the retrieved pictographs do not help to understand the meaning of the original sentence. 'Partially helpful' means that only one or several parts of the original meaning of the sentence can be understood on the basis of the pictographs that are retrieved, but not the meaning of the sentence as a whole. When we assess that the meaning of the entire sentence could be understood by looking at the pictographs, we use the category 'potentially fully helpful'. Considering the diversity of the intended users focused on in this paper, we deliberately use a tentative label. The evaluations were carried out by comparing the generated pictograph sequences to the original text. In order to establish inter-rater reliability, one independent annotator evaluated one third of the data set, following the same procedure. Despite the subjective nature of the evaluation procedure, Cohen's kappa (Cohen 1960) was 0.79, indicating that the agreement between both raters was (fairly) strong.

26. `https://osf.io/q2rmu/`

We also use four categories to analyse lexical coverage: ‘no pictograph’, ‘not clear/misleading’, ‘neutral’ and ‘potentially helpful’. We use the ‘neutral’ category when pictographs are neither clearly misleading nor potentially helpful. For this part of the evaluation, inter-rater reliability was higher than for the sentence-level analysis (Cohen’s kappa = 0.84).[27]

The same categories are used for evaluating the images retrieved from Wikipedia. Concrete examples of evaluations can be found in section 4.3.4.

4.3 Results

4.3.1 Sentence comprehensibility

Table 2 shows the results of the sentence comprehensibility analysis for the three contexts and the three pictograph sets. It is clear that the results for the language classes are better than for the newspaper articles and the administrative texts. *Sclera* and *Beta* obtain highly comparable results (with a slight advantage for *Sclera*), while both clearly outperform *Rand*. For three of the sentences in the administration data set the system timed out[28] for *Sclera* and *Beta*, resulting in no pictographs being retrieved. All three sentences consisted of more than 20 words.

	Language classes			**News**			**Administration**		
	Sclera	*Beta*	*Rand*	*Sclera*	*Beta*	*Rand*	*Sclera*	*Beta*	*Rand*
No pictographs	3	7	40	0	3	47	10	10	30
Not helpful	10	13	0	50	43	13	10	17	7
Partially helpful	47	47	47	40	50	40	80	73	63
Potentially fully helpful	40	33	13	10	3	0	0	0	0

Table 2: Results for sentence comprehensibility (in %)

Looking at the results for sentences taken from beginner language classes, the full meaning of two out of five (*Sclera*) or one out of three (*Beta*) sentences can potentially be represented by the pictograph translations. The percentage drops to 13% for *Rand* pictographs. This result is encouraging, especially considering that for around half of the sentences a partially helpful translation is generated. On the other hand, rendering a pictograph translation of a sentence that allows to fully understand the meaning of the original sentence seems nearly impossible for newspaper articles with the translation system and pictograph sets at hand, and impossible for administrative texts. Note that the sentences in the newspaper data set were taken from articles written in easy-to-read Dutch. More positively, partially helpful pictograph translations were obtained for between 40 and 50% of sentences in the news data and between 63 and 80% in the administration data set.

All in all, the sentences retrieved from the websites of Dutch-speaking communes around Brussels turn out to be the most difficult to translate, which is not that surprising considering their complexity. The fact that more partially helpful translations were retrieved for the administration data than for the news data can, at least in part, also be explained by the longer sentences: more lexical units can potentially be coupled with helpful pictographs. There were also fewer named entities in the administrative sentences.

27. The inter-rater reliability analysis further revealed that, perhaps unsurprisingly, agreement was lowest for the ‘neutral’ category.
28. Since limits on processing time are imposed for practical reasons, time-outs can occur, for example when the search space becomes too large.

4.3.2 Lexical coverage

The results of the lexical coverage analysis are presented in Table 3.[29] Here, the gap between the language classes on the one hand and the newspaper articles and administrative texts on the other, is smaller.

	Language classes			**News**			**Administration**		
	Sclera	*Beta*	*Rand*	*Sclera*	*Beta*	*Rand*	*Sclera*	*Beta*	*Rand*
No pictograph	29	26	79	27	30	85	41	42	85
Not clear/misleading	13	21	1	35	35	4	25	25	2
Neutral	5	5	4	5	3	2	5	6	2
Potentially helpful	52	49	16	32	32	8	29	28	10

Table 3: Results for lexical coverage (in %)

Around half of the lexical units are covered by a potentially helpful pictograph in the language classes data when using *Beta* and *Sclera*. For news and administration data this drops to around 30%. With *Rand* pictographs the coverage is markedly lower. In this context it is also important to note that, whereas for the language classes data potentially helpful pictographs clearly outnumber those that are deemed misleading, this is not the case for the news and administration data. For these data sets, retrieved pictographs are almost equally likely to be misleading or helpful.

The percentage of not covered lexical units is highest in the administrative data set. This can be explained by the specificity and low frequency of the vocabulary that is used here, coupled with the mentioned missing translation for three sentences.

For reference, named entities represented 7% of the lexical units in the language classes subset, 8% in the news data, and 4% in the administrative data set. None of these lexical units were covered by pictographs.

4.3.3 Retrieved images

Table 4 shows for how many named entities a picture was retrieved from Wikipedia using two different similarity thresholds, and whether this picture was deemed to be misleading, neutral or potentially helpful. Images were also retrieved for other lexical items, which are evaluated in Table 5.

	Threshold	
	1	*0.95*
Not covered	27	67
Not clear/misleading	15	3
Neutral	9	3
Potentially helpful	48	27

Table 4: Named entities (n=33) covered by images (in %)

Generally speaking, setting the parameter to 1 results in higher recall and lower quality, whereas setting it to 0.95 favours quality at the cost of coverage. With the threshold at 1, a potentially helpful image is retrieved for almost half of the named entities. In 15% of the cases the image retrieved is misleading. Conversely, with the threshold set to 0.95, the percentage of named entities for which a helpful pictograph is retrieved drops to 27%, but also the percentage of misleading images is reduced to 3%.

Table 5 shows that with the less strict threshold, 113 images were retrieved, compared to only 47 with threshold 0.95. One in three is deemed to be potentially helpful with threshold 1. This

29. These results concern lexical tokens, and not types, considering that the lexical units occur in different contexts, and potentially with different meanings.

	Threshold	
	1 (n=113)	*0.95 (n=47)*
Not clear/misleading	53	36
Neutral	17	23
Potentially helpful	30	40

Table 5: Results for other images added (in %)

increases to 40% with the stricter threshold. In contrast, the percentage of misleading images drops from 53% with threshold 1 to only 36% with the threshold fixed at 0.95.

4.3.4 Examples

In this section, we show a number of example pictograph translations to illustrate the results. These examples are selected to represent different evaluation categories.

Figure 7 shows the translation of the short sentence *Je koopt de krant* ('You buy a newspaper'), taken from the language classes data set, in *Sclera*, *Beta* and *Rand* pictographs. The *Sclera* and *Beta* translations were deemed to be potentially fully helpful, the translation into *Rand* pictographs only partially. In terms of lexical coverage, all three lexical units (i.e., *jij*, *koopt* and *krant*) were covered by *Sclera* and *Beta* pictographs, compared to only one for *Rand* pictographs. This sentence did not contain any named entities. No images were retrieved from Wikipedia.

Figure 7: *Je koopt de krant* ('You buy a newspaper') in *Sclera*, *Beta* and *Rand* pictographs

The translation into *Beta* pictographs of the sentence *In het weekend ga je een dag naar zee in Oostende* ('During the weekend you will spend a day at the sea in Ostend') is shown in Figure 8. This sentence is also taken from the language classes data set. The translation was rated as partially helpful. Note that the syntactic simplification module moved the temporal expression of time (*In het weekend*) to the end of the sentence, and switched the order of the subject and the verb to produce an S-V-O structure. After simplification, the sentence looked as follows: *Je gaat naar zee in Oostende een dag in het weekend.* The named entity *Oostende* is not covered by a pictograph. Note that the prepositions are also translated (rather literally), which may or may not be useful for interpretation.

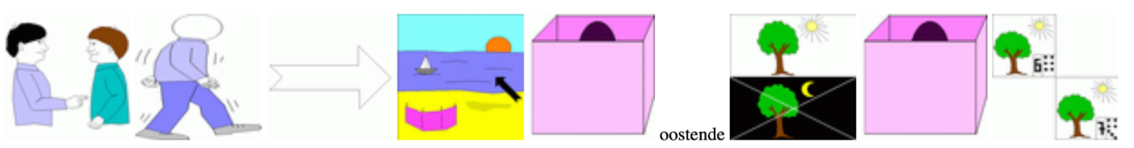

Figure 8: *In het weekend ga je een dag naar zee in Oostende* ('During the weekend you will spend a day at the sea in Ostend') in *Beta* pictographs

This input sentence contained one named entity (*Oostende*). Figure 9 shows the images that were retrieved from Wikipedia for this sentence, linked to the words *zee*[30] and *Oostende*.[31] These two images were linked to the input sentence when the similarity threshold was set to 1. With the stricter threshold, only the image representing *Oostende* was retrieved. Both images were deemed potentially useful. Whereas the input sentence is not specific about which sea is meant, the illustration shows the North Sea, which is the sea that is implied by the sentence. The image does contain text, which is not ideal, and for a correct interpretation some familiarity with the map of Northwestern Europe is probably necessary. In any case, a pictograph representing the word *zee* was also retrieved. The image retrieved for *Oostende* does represent this particular city, even though it may be difficult (for most people) to identify which city is shown exactly. In any case, it is clear that a coastal city is being referenced.

Figure 9: Retrieved Wikipedia images linked to *zee* ('sea', left) and *Oostende* ('Ostend', right)

Next, we look at the translation of the sentence *Ze werd bekend met het project 'Let's Go Urban' in Antwerpen* ('She became known with the project 'Let's Go Urban' in Antwerp.') into *Sclera* pictographs (Figure 10). This sentence is taken from the news corpus, and contains two named entities (*Let's Go Urban* and *Antwerpen*). It was part of an article that deals with a (relatively small-scale) political scandal in Flanders, that received a considerable amount of attention in the Flemish media. The syntactic simplification module identified an appositive clause (i.e., a sequential combination of nouns and/or noun phrases, *het project 'Let's Go Urban'*) and extracted this from the main clause, resulting in the following two sentences: *Ze werd bekend met het project in Antwerpen. Het project is 'Let's Go Urban'.*

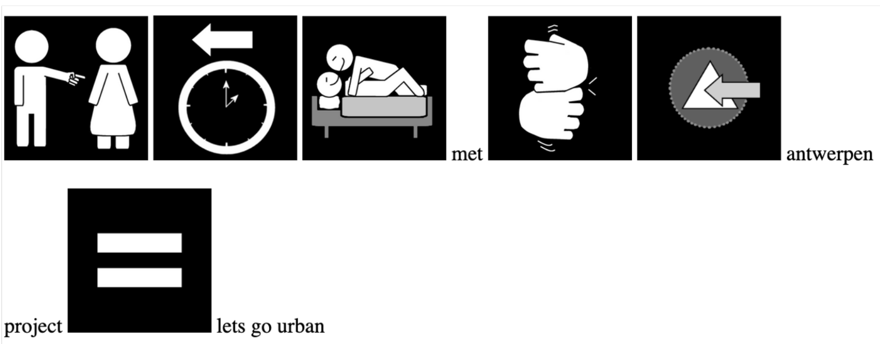

Figure 10: *Ze werd bekend met het project 'Let's Go Urban' in Antwerpen* ('She became known with the project 'Let's Go Urban' in Antwerp') in *Sclera* pictographs

30. Edited by M.Minderhoud - own work based on PD map. `https://upload.wikimedia.org/wikipedia/commons/a/ae/Locatie_Noordzee.PNG`
31. Marc Ryckaert - own work. `https://upload.wikimedia.org/wikipedia/commons/6/65/Oostende_Europacentrum_01.jpg`

The overall pictograph translation was rated as not helpful. Whereas the translation of *Je koopt de krant* shown in Figure 7 represented a best-case scenario, this pictograph sequence is probably our best example of how translation into pictographs can go wrong, and lead to highly confusing results. *Bekennen* is a (not widely used, archaic[32]) euphemism for sexual intercourse, similar to the English 'knowing'. Word sense disambiguation did not prevent the pictograph for *vrijen* ('make love') from being displayed for this input sentence. The combination with the pictograph for *werken* ('work') to its right is not conducive to a correct interpretation either.

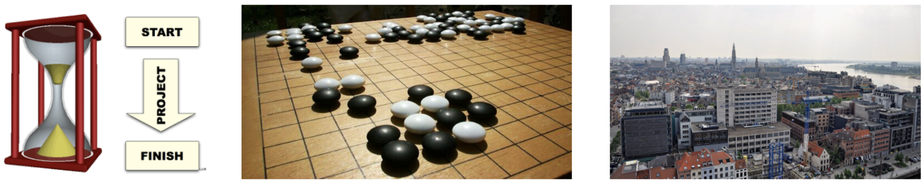

Figure 11: Retrieved Wikipedia images linked to *project*, *Go* and *Antwerpen*

Figure 11 lists the Wikipedia images that were linked to this sentence, and more specifically to the words *project*[33], *Go*[34] and *Antwerpen*.[35] These three images were only retrieved with the threshold fixed to 1. The first image is an artistic representation of a project, that does contain (English) text. In the context of this sentence, we classified it as 'not clear/misleading'. The image linked to the word 'Go' was also rated as misleading, as it shows the board game that goes by this name. Note that we had identified one named entity consisting of multiple words (i.e., *het project 'Let's Go Urban'*), for which no relevant Wikipedia page was retrieved. The third image is a photograph of Antwerp, so we classified it as potentially helpful, with the same caveat we mentioned for the picture of Ostend.

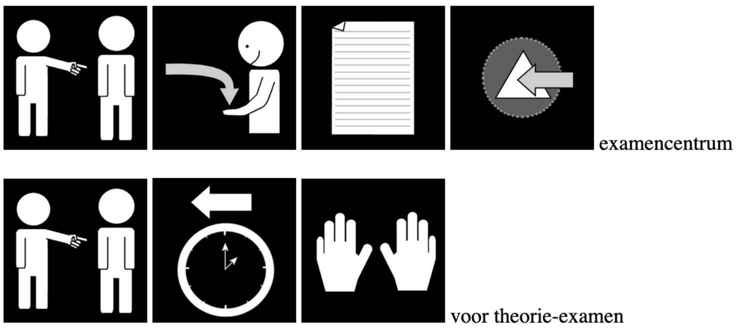

Figure 12: *U krijgt een aanvraagformulier in het examencentrum als u voor het theorie-examen bent geslaagd* ('You receive an application form in the examination center if you pass the theoretical exam') in *Sclera* pictographs

Finally, we will consider the pictograph translations for a sentence taken from the administration data set. Figure 12 shows the *Sclera* translation of *U krijgt een aanvraagformulier in het examencen-*

32. See, for example, Genesis 4:1, *En Adam bekende Heva, zijn huisvrouw, en zij werd zwanger* (Statenbijbel, 1637), 'And Adam knew Eue his wife, and shee conceiued' (King James Bible, 1611).
33. Laurens van Lieshout, January 2006. `https://upload.wikimedia.org/wikipedia/commons/5/5c/Project_artistic_impression.png`
34. Donarreiskoffer - own picture. `https://upload.wikimedia.org/wikipedia/commons/2/2e/Go_board.jpg`
35. Paul Hermans - own work `https://upload.wikimedia.org/wikipedia/commons/f/f0/Stadsgezicht_van_Antwerpen_vanaf_het_MAS_30-05-2012_15-29-35.jpg`

trum als u voor het theorie-examen bent geslaagd ('You receive an application form in the examination center if you pass the theoretical exam'), which was taken from a local government's web page with information on how to obtain a driving licence. Following syntactic simplification, the sentence was split in two: *U krijgt een aanvraagformulier in het examencentrum. U bent geslaagd voor het theorie-examen.* We rated this translation as potentially partially helpful, since the first three pictographs represent the first part of the sentence rather well. The meaning of the if-clause, however, is entirely lost. Note that no pictographs were retrieved for the compound nouns *examencentrum* ('examination center') and *theorie-examen* ('theoretical exam'), whereas for *aanvraagformulier* ('application form') the pictograph linked to the hyperonym *papier* ('(sheet of) paper') was produced. This sentence did not contain named entities, and no Wikipedia images were retrieved.

The translation into *Beta* pictographs of the same sentence is shown in Figure 13. Since in this case no pictograph was retrieved for the object *aanvraagformulier*, we deemed the translation to be not helpful.

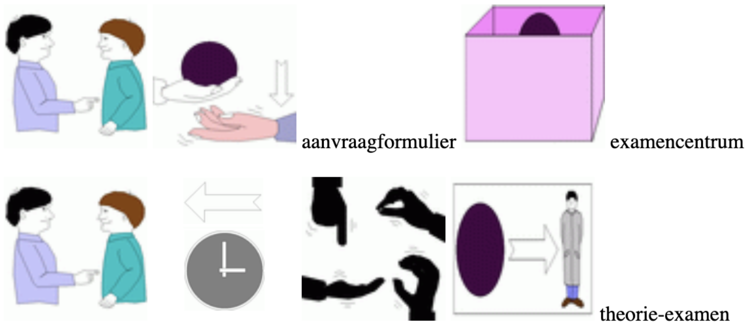

Figure 13: *U krijgt een aanvraagformulier in het examencentrum als u voor het theorie-examen bent geslaagd* in *Beta* pictographs

For the translation into *Rand* pictographs, only one image was retrieved (Figure 14). In spite of this poor lexical coverage, it could be argued that this single pictograph represents the if-clause fairly well, so therefore we rated the translation of the sentence as potentially partially helpful.

Figure 14: *U krijgt een aanvraagformulier in het examencentrum als u voor het theorie-examen bent geslaagd* in *Rand* pictographs

5. Discussion

Our hypothesis stated that the task at hand would be difficult, and this is confirmed by the experiments. Especially the results for overall sentence comprehensibility (RQ1) show that the generated pictograph translations are unable to convey the full meaning of the input sentences. This is true in particular for sentences taken from newspaper articles and websites of local governments. The results for the beginner language classes data set are more encouraging, even though the meaning of well over half of the sentences was not fully represented by the pictographs. This leads us to conclude that the text-to-pictograph system tested here is, in its current state of development, not suited to be used directly by migrants themselves. The fact that certain pictograph translations can clearly lead to confusion, as some of our examples illustrate, strengthens this conclusion. Considering also that some potential contexts of use are high-stakes and mistakes can be costly, we do not think the current system is accurate enough to be used as such in the contexts investigated in this study.

The results for lexical coverage (RQ2) indicate that, even though it is difficult for the translation system to preserve the full meaning of sentences, it can successfully generate pictographs that represent a considerable number of lexical units. Unsurprisingly, the more frequent and concrete a word is, the more likely it is to be coupled with a relevant pictograph. Lexical simplification may be an option to extend the coverage of the system (Bulté et al. 2018), even though this is already partially covered by the synonym sets that are used. Our experiments also show that the coverage is much lower when *Rand* pictographs are used, compared to *Sclera* and *Beta*. It would be interesting to test the coverage of other pictograph sets, such as *ARASAAC*, which has recently been linked to a French version of the text-to-pictograph translation tool (Norré et al. 2021).

Our tests with image retrieval from Wikipedia (RQs 3 and 4) suggest that this is a viable option to increase the coverage of the system. However, not all relevant Wikipedia articles that are retrieved contain images, and some of the images that are retrieved are more likely to lead to confusion than to aid comprehension. To increase the coverage of the system, it could be an option to also include Wikipedia pages in languages other than Dutch. Since image retrieval seems most successful for named entities, performing named entity recognition first might be a useful intermediate step.

Although the performance of the system is far from optimal, the results of our evaluations have to be considered in the light of the complexity of the task at hand. For a start, it would already be difficult to manually find or design good pictographs to cover many of the individual concepts and complex meanings expressed in the three contexts, and the system is also not developed for a particular domain. On top of this, the end users have different cultural and language backgrounds, and they are not trained to communicate with pictographs, in contrast to the originally intended end-users. This being said, we do believe that text-to-pictograph conversion can be useful for, amongst others, language teachers, caregivers, government officials and social workers who wish to augment written text with pictographs and/or images. With this purpose in mind, we intend to further adapt the translation system so that it, for example, offers alternatives from which a content creator can easily choose the best fitting ones for a particular context and message they want to convey.

This study has a number of limitations, especially related to the evaluation procedure. The data sets used for testing were fairly small, consisting of 30 sentences each. Not only were the evaluations to some extent subjective, they were carried out by the researchers themselves and not by the intended end-users. It would be necessary to conduct tests with, for example, language teachers, to confirm our tentative conclusions about the usefulness of the tool for applications with a human-in-the-loop.

6. Conclusion

We set out to evaluate the potential of a text-to-pictograph translation system in a Flemish migration setting, considering that pictographs are often used as a tool to facilitate communication in such a context. Pictographs are not only used in communication directed at migrants or in (beginner) language classes, they are also made available to migrants in the form of booklets and online searchable databases to assist them in their daily lives. Such tools are not always easy or fast to use, so an automatic text-to-pictograph translation system has the potential to be a useful addition. We tested the potential usefulness of pictograph translations of sentences taken from three authentic contexts (beginner language classes, news articles, administrative communication), using three different sets of pictographs. These tests were carried out without a 'human-in-the-loop' in mind, meaning that the pictograph translations were evaluated from the point of view of potential users that are unaware of the meaning of the original Dutch sentence. In our hypotheses, we acknowledged that this would be a challenging task, for a number of interconnected reasons, and our results confirm this. We concluded that the text-to-pictograph translation system should not be used by migrants as such. It could be used to obtain 'gist' translations, but this would mainly be useful for simple, preferably short sentences that do not contain specific and infrequent vocabulary.

To end on a more positive note, our results are promising with other potential use cases in mind. For example, a pictograph translation system could be useful in educational settings, such as language classes for beginners with mixed language backgrounds and varying degrees of literacy. Teachers could use the tool to create content for their classes with added visual support. The system could also be useful in general to facilitate communication with migrants, in situations where generated translations are verified and/or when they are used to create pictograph-based communications.

Acknowledgements

We would like to thank the Humanities and Social Sciences Group at KU Leuven, which allowed us to further develop and valorise the pictograph translation tools within the context of the 2018 Valorisation Fund. We also thank the staff of vzw De Rand for their valuable feedback and the integration of our tools on their website. Finally, we would like to thank the reviewers and editors for their valuable comments.

References

Agrawal, Rakesh, Sreenivas Gollapudi, Anitha Kannan, and Krishnaram Kenthapadi (2011), Enriching textbooks with images, *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*, pp. 1847–1856.

Auer, Daniel (2018), Language roulette–the effect of random placement on refugees' labour market integration, *Journal of Ethnic and Migration Studies* **44** (3), pp. 341–362.

Barros, Izadora M.C., Thaciana S. Alcântara, Alessandra R. Mesquita, Anne Caroline O. Santos, Felipe P. Paixão, and Divaldo P. Lyra Jr (2014), The use of pictograms in the health care: A literature review, *Research in Social and Administrative Pharmacy* **10** (5), pp. 704–719.

Bautista, Susana, Raquel Hervás, Agustín Hernández-Gil, Carlos Martínez-Díaz, Sergio Pascua, and Pablo Gervás (2017), Aratraductor: Text to pictogram translation using natural language processing techniques, *Proceedings of the XVIII International Conference on Human Computer Interaction*, Association for Computing Machinery.

Böcker, Martin (1996), A multiple index approach for the evaluation of pictograms and icons, *Computer Standards and Interfaces* **18** (2), pp. 107–115.

Brank, Janez, Gregor Leban, and Marko Grobelnik (2017), Annotating documents with relevant Wikipedia concepts, *Proceedings of the Slovenian Conference on Data Mining and Data Warehouses.*

Bulté, Bram, Leen Sevens, and Vincent Vandeghinste (2018), Automating lexical simplification in Dutch, *Computational Linguistics in the Netherlands Journal* **8**, pp. 24–48.

Caffaro, Federica and Eugenio Cavallo (2015), Comprehension of safety pictograms affixed to agricultural machinery: A survey of users, *Journal of Safety Research* **55**, pp. 151–158.

Cohen, J. (1960), A Coefficient of Agreement for Nominal Scales, *Educational and Psychological Measurement* **20** (1), pp. 37–46.

Cowgill, Jamie and Jim Bolek (2003), *Symbol usage in health care settings for people with limited English proficiency. Part One - Evaluation of Use of Symbol Graphics in Medical Settings*, JRC Design, Scottsdale, AZ.

Cox, Antoon (2017), *The dynamics of (mis)communication in language discordant multi-party consultations in the Emergency Department*, Dissertation, University Press, Zelzate.

Coyne, Bob and Richard Sproat (2001), Wordseye: An automatic text-to-scene conversion system, *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 487–496.

Davies, Sarah, Helen Haines, Beverley Norris, and John R. Wilson (1998), Safety pictograms: are they getting the message across?, *Applied Ergonomics* **29** (1), pp. 15–23.

Dowse, Roslind and Martina S Ehlers (1998), Pictograms in pharmacy, *International Journal of Pharmacy Practice* **6** (2), pp. 109–118.

Fedasil (2019), *Annual Report 2019*, Brussels, Belgium.

Goonetilleke, Ravindra S, Heloisa Martins Shih, and Julien Fritsch (2001), Effects of training and representational characteristics in icon design, *International Journal of Human-Computer Studies* **55** (5), pp. 741–760.

Hamann, Ulrike and Nihad El-Kayed (2018), Refugees' access to housing and residency in German cities: internal border regimes and their local variations, *Social Inclusion* **6** (1), pp. 135–146.

Hooft, Hannelore, Hanne Vandermeerschen, Saena Chakkar, Goedele Vandommele, Peter De Cuyper, and Mariet Schiepers (2020), *Laaggeletterdheid bij volwassen nieuwkomers in de Belgische opvang. Een verkennende analyse*, Centrum voor Taal en Onderwijs & Instituut voor Arbeid en Samenleving, Leuven, Belgium.

Huer, Mary B (2000), Examining perceptions of graphic symbols across cultures: Preliminary study of the impact of culture/ethnicity, *Augmentative and Alternative Communication* **16** (3), pp. 180–185.

IOM (2019), *World Migration Report 2020*, International Organization for Migration, Geneva, Switzerland.

Jensen, Natasja Koitzsch, Marie Norredam, Stefan Priebe, and Allan Krasnik (2013), How do general practitioners experience providing care to refugees with mental health problems? A qualitative study from Denmark, *BMC Family Practice* **14** (1), pp. 1–9.

Joshi, Dhiraj, James Z Wang, and Jia Li (2006), The story picturing engine—a system for automatic text illustration, *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* **2** (1), pp. 68–89.

Keskinen, Tuuli, Tomi Heimonen, Markku Turunen, Juha-Pekka Rajaniemi, and Sami Kauppinen (2012), Symbolchat: A flexible picture-based communication platform for users with intellectual disabilities, *Interacting with Computers* **24** (5), pp. 374–386.

Koehler, Claudia and Jens Schneider (2019), Young refugees in education: the particular challenges of school systems in Europe, *Comparative Migration Studies* **7** (1), pp. 1–20.

Krystallidou, Demi, Wolf Langewitz, and Maria van den Muijsenbergh (2021), Multilingual healthcare communication: Stumbling blocks, solutions, recommendations, *Patient Education and Counseling* **104** (3), pp. 512–516.

Leemans, Paul (2001), *Vil: a visual inter lingua*, Dissertation, Worcester Polytechnic Institute.

Lochmann, Alexia, Hillel Rapoport, and Biagio Speciale (2019), The effect of language training on immigrants' economic integration: Empirical evidence from France, *European Economic Review* **113**, pp. 265–296.

Manirankunda, Lazare, Jasna Loos, Pieterjan Debackaere, and Christiana Nöstlinger (2012), "It is not easy": challenges for provider-initiated HIV testing and counseling in Flanders, Belgium, *AIDS Education and Prevention* **24** (5), pp. 456–468.

Marrero, Mónica, Julián Urbano, Sonia Sánchez-Cuadrado, Jorge Morato, and Juan Miguel Gómez-Berbís (2013), Named entity recognition: fallacies, challenges and opportunities, *Computer Standards and Interfaces* **35** (5), pp. 482–489.

Martín, Alejandro, Raquel Hervás, Gonzalo Méndez, and Susana Bautista (2018), *PICTAR: Una herramienta de elaboración de contenido para personas con TEA basada en la traducción de texto a pictogramas*, MA thesis, Universidad Complutense de Madrid.

Martínez-Santiago, Fernando, Manuel Carlos Díaz-Galiano, Luis Alfonso Ureña-López, and Ruslan Mitkov (2015), A semantic grammar for beginning communicators, *Knowledge-Based Systems* **86**, pp. 158–172.

McCoy, Kathleen F (1998), Interface and language issues in intelligent systems for people with disabilities, *in* Mittal, Vibhu O., Holly A. Yanco, John Aronis, and Richard C. Simpson, editors, *Assistive Technology and Artificial Intelligence*, Springer, pp. 1–11.

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

Miller, George A (1995), WordNet: a lexical database for English, *Communications of the ACM* **38** (11), pp. 39–41.

Montagne, Michael (2013), Pharmaceutical pictograms: a model for development and testing for comprehension and utility, *Research in Social and Administrative Pharmacy* **9** (5), pp. 609–620.

Müller, Ingrid, Margret Buchholz, and Ulrika Ferm (2010), Text messaging with picture symbols - experiences of seven persons with cognitive and communicative disabilities, *Journal of Assistive Technologies* **4** (4), pp. 11–23.

Munemori, Jun, Taro Fukuda, Moonyati Binti Mohd Yatid, Tadashi Nishide, and Junko Itou (2010), Pictograph chat communicator III: a chat system that embodies cross-cultural communication, *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, pp. 473–482.

Nadeau, David and Satoshi Sekine (2007), A survey of named entity recognition and classification, *Lingvisticae Investigationes* **30** (1), pp. 3–26.

Norré, Magali, Pierrette Bouillon, Johanna Gerlach, and Hervé Spechbach (2021), Evaluating the comprehension of Arasaac and Sclera pictographs for the BabelDr patient response interface, *Proceedings of the 3rd Swiss conference on barrier-free communication (BfC 2020)*, pp. 55–63.

Norré, Magali, Vincent Vandeghinste, Pierette Bouillon, and Thomas François (2021), Extending a text-to-pictograph system to French and to Arasaac, *International Conference Recent Advances in Natural Language Processing (RANLP).*

Paolieri, Daniela and Alejandra Marful (2018), Norms for a pictographic system: the Aragonese portal of augmentative/alternative communication (ARASAAC) system, *Frontiers in Psychology* **9**, pp. 2538.

Pavli, Androula and Helena Maltezou (2017), Health problems of newly arrived migrants and refugees in Europe, *Journal of Travel Medicine* **24** (4), pp. 1–8.

Poots, J Kent and Ebrahim Bagheri (2018), Automatic annotation of text with pictures, *IT Professional* **20** (1), pp. 36–44.

Postma, Marten, Emiel van Miltenburg, Roxane Segers, Anneleen Schoen, and Piek Vossen (2016), Open Dutch WordNet, *Proceedings of the Eight Global Wordnet Conference*, Bucharest, Romania.

Romski, MaryAnn, R Sevcik, A Smith, RM Barker, Stephanie Folan, and Andrea Barton-Hulsey (2009), The system for augmenting language: Implications for young children with autism spectrum disorders, *in* Mirenda, Pat and Teresa Iacono, editors, *Autism Spectrum Disorders and AAC*, Paul H. Brookes, Baltimore, MD, pp. 219–245.

Schuurman, Ineke, Vincent Vandeghinste, and Leen Sevens (2020), Wordnet, occupations and natural gender, *Presented at the 30th conference for Computational Linguistics in the Netherlands.*

Sevens, Leen (2018), *Words Divide, Pictographs Unite: Pictograph Communication Technologies for People with an Intellectual Disability*, Dissertation, LOT.

Sevens, Leen, Gilles Jacobs, Vincent Vandeghinste, Ineke Schuurman, and Frank Van Eynde (2016), Improving text-to-pictograph translation through word sense disambiguation, *Proceedings of the Fifth Joint Conference on Lexical and Computational Semantics*, Berlin, Germany, pp. 131–135.

Sevens, Leen, Vincent Vandeghinste, Ineke Schuurman, and Frank Van Eynde (2015), Extending a Dutch text-to-pictograph converter to English and Spanish, *Proceedings of SLPAT 2015: 6th Workshop on Speech and Language Processing for Assistive Technologies*, pp. 110–117.

Sevens, Leen, Vincent Vandeghinste, Ineke Schuurman, and Frank Van Eynde (2018), Involving People with an Intellectual Disability in the Development of Pictograph Translation Technologies for Social Media Use, *in* Cougnon, Louise-Amélie, editor, *Language and the new (instant) media*, Presses Universitaires de Louvain, Louvain-la-Neuve, pp. 57 – 68.

Sinclair, John (1998), The lexical item, *in* Weigand, Edda, editor, *Contrastive Lexical Semantics*, John Benjamins, pp. 1–24.

Soares, Marcos André Barroso (2015), *Designing Culturally Sensitive Icons for User Interfaces: An approach for the Interaction Design of smartphones in developing countries*, Dissertation, Universidade do Porto.

Stevenson, Jacqueline and Sally Baker (2018), *Refugees in Higher Education: Debate, Discourse and Practice*, Emerald Group Publishing.

Tanimoto, Steven L (1997), Representation and learnability in visual languages for web-based interpersonal communication, *Proceedings. 1997 IEEE Symposium on Visual Languages (Cat. No. 97TB100180)*, pp. 2–10.

van Noord, Gertjan (2006), At last parsing is now operational, *TALN 2006*, pp. 20–42.

Vandeghinste, Vincent, Ineke Schuurman, Leen Sevens, and Frank Van Eynde (2017), Translating text into pictographs, *Natural Language Engineering* **23** (2), pp. 217–244.

Vandeghinste, Vincent (2002), Lexicon optimization: Maximizing lexical coverage in speech recognition through automated compounding, *Proceedings of the Third International Confererence on Language Resources and Evaluation (LREC)*, pp. 1270–1276.

Vandeghinste, Vincent and Ineke Schuurman (2014), Linking pictographs to synsets: Sclera2Cornetto, *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland, pp. 3404–3410.

Vandeghinste, Vincent, Bram Bulté, and Liesbeth Augustinus (2019), Wablieft: An easy-to-read newspaper corpus for Dutch, *Proceedings of CLARIN Annual Conference 2019*, pp. 188–191.

Vossen, Piek T. J. M., Isa Maks, Roxane Segers, Hennie van der Vliet, Marie-Francine Moens, Katja Hofmann, Erik Tjong Kim Sang, and Maarten de Rijke (2013), Cornetto: a combinatorial lexical semantic database for Dutch, *in* Spyns, Peter and Jan Odijk, editors, *Essential Speech and Language Technology for Dutch. Results by the STEVIN-programme*, Springer, pp. 165–184.

Wauters, Bram and Johan Lambrecht (2008), Barriers to refugee entrepreneurship in Belgium: Towards an explanatory model, *Journal of Ethnic and Migration Studies* **34** (6), pp. 895–915.

Appendix A. Data sets used for evaluation

Document	*Sample sentences*
Language classes	
Niveau 1.1. lesblok 1: Jezelf voorstellen	• Anna en Noah starten met een opleiding in een avondschool. • Ze doen een graduaat marketing.
Niveau 1.1. Lesactiviteit: "De trein nemen (3)"	• Je woont in Anderlecht. • Deze avond heb je een afspraak met een vriend.
Niveau 1.1. Lesactiviteit: "Zoekertjes"	• Je koopt de krant. • In de krant staan zoekertjes.
Niveau 1.1. Aan het werk	• Je werkt al een jaar bij een groot bedrijf in België. • Je bent niet zo tevreden met je werksituatie.
News	
Weer strenger tegen virus (24.03.2021)	• Corona wordt te gevaarlijk. • Het virus blijft groeien.
Grote twijfels over El Kaouakibi (02.04.2021)	• Open Vld wil politica Sihame El Kaouakibi voorgoed uit de partij zetten. • Dat gebeurt na veel verhalen over fraude.
Afschuw over geweld in Myanmar (31.03.2021)	• De toestand in Myanmar wordt steeds erger. • Het leger van dat land greep twee maanden geleden de macht.
Administration	
Welzijn en OCMW - Leefloon (Dilbeek)	• Het leefloon is een onderdeel van het recht op maatschappelijke integratie. • Hebt u geen inkomen of is dit lager dan bepaalde grensbedragen, dan ontvangt u onder bepaalde voorwaarden van het OCMW een leefloon.
Voorlopig rijbewijs aanvragen (Vilvoorde)	• Een voorlopig rijbewijs geeft u de toelating om met een voertuig op de openbare weg te rijden als voorbereiding op het praktisch rijexamen. • U krijgt een aanvraagformulier in het examencentrum als u voor het theorie-examen bent geslaagd.
Taaltest en advies (Zaventem)	• Indien je je wil inschrijven in een CVO (centrum voor volwassenonderwijs) of CBE (Centrum voor Basiseducatie) heb je van ons een attest nodig. • Tijdens een intakegesprek stellen we daarom vragen over je verwachtingen, talenkennis, hoogst behaalde diploma,...

Table A1: Overview data sets with sample sentences