Translating a low-resource language using GPT-3 and a human-readable dictionary

Micha Elsner and **Jordan Needle**
Department of Linguistics
The Ohio State University
Columbus, Ohio
{elsner.14,needle.6}@osu.edu

Abstract

We investigate how well words in the polysynthetic language Inuktitut can be translated by combining dictionary definitions, without use of a neural machine translation model trained on parallel text. Such a translation system would allow natural language technology to benefit from resources designed for community use in a language revitalization or education program, rather than requiring a separate parallel corpus. We show that the text-to-text generation capabilities of GPT-3 allow it to perform this task with BLEU scores of up to 18.5. We investigate prompting GPT-3 to provide multiple translations, which can help slightly, and providing it with grammar information, which is mostly ineffective. Finally, we test GPT-3's ability to derive morpheme definitions from whole-word translations, but find this process is prone to errors including hallucinations.

1 Introduction

In low-resource language communities, resource creation efforts are restricted by the limited time community members can contribute— and this problem is worsened when effort must be divided between development of community-facing resources and those targeted at machines. Language revitalization and pedagogy programs need dictionaries (especially those which incorporate tools for morphological analysis and flexible search) and grammar lessons, while machine translation systems need large corpora. If community-facing resources could be used within machine learning systems to compensate for the limited availability of text, community efforts could serve both pedagogical and technological goals at the same time. But while this has occasionally been attempted, techniques for doing so are still not effective enough to serve as standard methods.

One reason why community-facing resources like bilingual dictionaries have not been widely used in applications like low-resource translation is that understanding definitions can require sophisticated tools for natural language understanding. Recent advances in large language model technology (LLMs) provide a promising candidate for such a toolset, at least for definitions written in high-resource languages. In this paper, we investigate methods for using a representative LLM, GPT-3, to translate multi-morphemic words in Inuktitut using dictionary definitions for the morphemes. In addition, we measure GPT-3's capacity to perform the reverse task of inferring the dictionary definition of a morpheme given a decomposed word in which it occurs.

Inuktitut is a polysynthetic language in which words can be very long and morphologically complex. As such, it is representative of a number of languages of the Americas for which natural language processing tasks have historically been difficult due to limited resources and typological differences from better-resourced languages (Mager et al., 2018a). Computational tools are an important part of education or revitalization efforts for American languages, including Inuktitut (Ngoc Le and Sadat, 2020).

We show that GPT-3 can stitch together English dictionary definitions to produce reasonable translations of many Inuktitut words. We investigate two further questions: methods for dealing with morphemes with multiple definitions, and the extent to which performance can be improved by priming GPT-3 with some grammatical information. We envision our system as one component of an interactive dictionary/translation system, in which a human learner or non-native speaker asks for possible analyses of a morphologically complex form and is given both a morph-by-morph gloss and some possible translations into fluent English. This kind of system might help to bridge the gap between a conventional dictionary, which is incapable of interactively translating morphologically complex words into fluent English, and full-scale neural ma-

Proceedings of the 20th SIGMORPHON workshop on Computational Research in Phonetics, Phonology, and Morphology, pages 1–13
July 14, 2023

chine translation (NMT), which is data-hungry and non-transparent. We further see possibilities for suggesting new dictionary definitions which can be curated by native speakers. Finally, we believe it holds some potential as a stepping stone towards larger-scale NMT applications by providing simple examples for use in a curriculum learning paradigm (Platanios et al., 2019; Liang et al., 2021, among others).

2 Related work

The use of bilingual dictionary entries in neural machine translation was pioneered by Luong et al. (2015), who augment an English/French MT system with mechanisms for aligning rare words across languages, then copying material from definitions to translate them. This work makes several key assumptions about the benefits dictionary definitions can provide. In particular, it assumes a relatively capable NMT system already exists. Because of this, the main contribution of the dictionary is to provide lexical equivalents for rare items. In most cases, these are content words, and their definitions, once known, are easily integrated into the translated sentence (Dinu et al., 2019; Zhang et al., 2021). Pham et al. (2018); Niehues (2021) are among the earliest to provide dictionary information as augmented input rather than via a custom architecture (thus moving toward a zero-shot method), but still requires the NMT system to be trained to use definitions.

Our work follows from recent approaches which use LLMs rather than purpose-built NMT systems. The ability of LLMs to translate some high-resource language pairs in a prompt-based zero or few-shot setting is established by Brown et al. (2020). Some other recent papers attempt to augment LLMs with information derived from dictionaries or phrase tables. Sun et al. (2022) translates full sentences, using hints from a phrase aligner (Dou and Neubig, 2021). Their approach also emphasizes the monolingual text-to-text generation capacity of LLMs, but uses aligned phrases rather than dictionary definitions intended for human readers. The closest point of comparison to our work is Ghazvininejad et al. (2023), which improves translation performance by augmenting few-shot prompts with dictionary translations for a few selected words. Unlike our setting, where the baseline NMT system performs poorly, their languages are selected so that the baseline NMT model performs reasonably without augmentation (10-30% BLEU).

In contrast to these approaches, we assume a setting in which common morphs, including functional as well as content items, must be translated with the aid of the dictionary. While this is an artificial constraint in the case of Inuktitut, which does have enough parallel text to train an NMT system, it is the case for other low-resource polysynthetic languages of the Americas which lack parallel corpora large enough to train any NMT system. Even where data is more plentiful, dictionary definitions are potentially helpful for translating functional items in polysynthetic languages because these languages can have very large paradigms with very unbalanced attestations in corpora. Inuktitut has a polypersonal agreement system in which subject and object person and number are both marked on verbs; some subject-object markers rarely appear in written corpora due to discourse constraints (for example, dual subjects). Other morphosemantic distinctions (such as dual number, evidentiality, intensifiers and applicatives) which are common in American languages but rare in European ones, are translated in very different ways across contexts, leading MT systems to misalign them due to limited data (Mager et al., 2018b).

In this setting, dictionary definitions may be patched together in relatively complex ways. First, composing the definitions is more difficult than simple concatenation: *takujara* "I see him/her" is made up of *taku* "see" and *jara* "I ... him/her" (1SG>3SG). Second, Inuktitut uses derivational processes to express terms which have independent content words in English. *qukiut* "gun" is made up of *qukit* "shoot" and the instrumental marker *ut*. While a literal translation would produce "an instrument for shooting", the leap to paraphrase this expression as "gun" requires a deeper representation of English semantics. Thus, while previous systems could use dictionary material mainly by copying, our task setting emphasizes text-to-text generation.

Related tasks which use language modeling to patch together fragments of target-language structure include bag-to-sequence word ordering (Hasler et al., 2017) and dependency linearization (Mille et al., 2020). Generating fluent text from grammatical element annotations is also similar to generating translations from glosses (Zhang and Duh, 2021; Garera and Yarowsky, 2008), although, despite recent interest in glosses (e.g. Moeller and Hulden, 2021), this task is also comparatively understudied.

Inuktitut itself is one of the best-resourced indigenous American languages, with a large parallel corpus collected from the Nunavut Parliamentary Hansards (Joanis et al., 2020); their baseline NMT system yields an IU→EN BLEU score of 35.0. This dataset was used as a challenge for the Workshop on Machine Translation in 2020 (Barrault et al., 2020). Scores remained low compared to better-resourced languages— the system rated highest by humans reports a BLEU score of only 29.1 on test (Zhang et al., 2020). As stated, we use Inuktitut as a potential model for less-resourced polysynthetic languages where even this level of NMT performance is not available.

In addition to assuming access to a dictionary, we also assume access to a system which provides canonical morphological segmentations (based on Farley (2012) and further described in Section 3), so that a complex word can be decomposed into parts whose lexical entries can be accessed. Such systems (Wiemerslage et al., 2022) can be developed with access to substantially less data than NMT systems. They may be created using finite-state toolkits (Park et al., 2021) or supervised learning from relatively small annotated datasets (Mager et al., 2020; Liu et al., 2021), potentially incorporating active learning (Grönroos et al., 2016).

Segmentation of Inuktitut is easier than translation— Uqailaut (Farley, 2012) is a widely used finite-state system for canonical segmentation of Inuktitut. Micher (2017) describes an improved segmenter with neural disambiguation; Roest et al. (2020) conduct more recent experiments on segmentation using Transformer networks. While our assumption that a segmentation system is available constitutes a weakness of our method, we hope that future work can continue to reduce the resource burden of developing such systems and can also tie their development more closely to community-facing resources such as grammar texts.

3 Data and metrics

We extract a lexicon of Inuktitut morphemes and their definitions from the Uqausiit dictionary (`uqausiit.ca`) created by Inuit Uqausinginnik Taiguusiliuqtiit, an Inuit language authority funded by the Nunavut Legislature. Uqausiit also provides example phrases with English translations and hand-annotated partial morphological decompositions. We extract all single-word example phrases with a multi-morphemic partial decomposition for development and testing of our translation systems. Tables 1, 2 and 3 show statistics of the Uqausiit datasets used in this paper.

Whole words	Count	With seg.
Dev.	50	22
Test	448	219

Table 1: Statistics of the translation data from Uqausiit.

Target morpheme	Instances	Unique
Root	219	89
Functional	218	130

Table 2: Statistics of the definition prediction data from Uqausiit.

Morphemes	Count
Root	2782
Grammatical	1524
Variable	157
Total	4462

Table 3: Statistics of the Uqausiit morpheme dictionary.

Because our aim is primarily to evaluate the potential of LLMs to operate on dictionary definitions, rather than to evaluate algorithms for segmentation, we use a fixed canonical morphological segmentation for each word in our dataset. These are provided by a partial oracle which is implementationally simple to create. We run the Uqailaut FST segmenter (Farley, 2012), which produces a set of candidate analyses. We then intersect these analyses with the partial decomposition of the phrase from Uqausiit. If one or more analyses match, we select the first one and return it. If Uqailaut cannot analyze the word, or produces no analyses matching the partial decomposition, we do not use the word. (Many of these errors result from orthographic or dialectal variation beyond the scope of Uqailaut's design, as noted by Mallon (2000).)

For instance, *sikujuittuq* "an area of the ocean where ice does not form" has the Uqausiit partial segmentation *siku-juit*, without the final *tuq*. Uqailaut produces 6 candidate segmentations, which vary as to the analysis of the medial *juitt* sequence; of these, we select *siku-juit-juq* since this matches the provided partial segmentation. (The resulting segmentation may still contain an incorrect element which is not part of the partial decomposition; we have no way to measure how often this occurs, but did not find cases in development.)

We also create datasets for creating dictionary definitions, for both root and functional morphemes. In each case, we use a whole word with its translation as the prompt, and query the definition of one of the component morphemes. We assume the root morpheme is the first one in the word; for functional morphemes, we select a later one at random. We create instances for all 219 segmentable words in the test set.[1] Some of these instances ask for a definition of the same query morpheme (but with different whole-word prompts); there are 89 unique root types and 130 unique functional types in the dataset. Where a query morpheme has multiple possible definitions, we refer to Uqausiit and the whole-word definition to select the correct one as a reference.

We evaluate translations against the English references using Sacrebleu (Post, 2018) and BLEURT (Sellam et al., 2020)[2]. In some cases, we ask the system to produce multiple candidate translations. We anticipate a human user of the dictionary considering the context in which they encountered the phrase they are looking up and picking the best one; this may be especially helpful in cases where the phrase is actually ambiguous in its translation, since otherwise there is no way to pick between candidate meanings. For these, we evaluate BLEURT in two ways: the **average** performance is the expected score of each response against the reference, and reflects the user's experience if they are looking up a word for which they have no useful context. The **oracle** score is the best score of any response, and reflects the user's experience if they can always pick the correct translation given the context in which they heard the target word.

Because Sacrebleu produces a global precision score over the whole corpus, we evaluate the expected Sacrebleu in multiple-translation cases by sampling one translation from the set of proposals for each word; we average across five samples. We produce an **oracle** Sacrebleu score by collating the translations of each word with the highest BLEURT score and evaluating them as a group.

[1] One functional item had no listed definition and was discarded.

[2] Using the recommended BLEURT-20 checkpoint.

The Inuktitut word saviggirunnaqtutit is made up of the following parts:
savik: (1) metal; steel; iron (2) snow-knife; or, to be fitted with a metal point (harpoon; spear)
ggiq: to bring someone or something along, as in "Don't bring your gun along."
runnaq: to be able to perform a certain action, as in "Could you find that out if he/she arrived"
tutit: you, as in "you sleep"
In English, saviggirunnaqtutit means roughly:

Figure 1: An example prompt for the **definition+example** case.

4 Single-word translation

All our experiments use the OpenAI API to access the GPT-3 model TEXT-DAVINCI-003, which was the largest GPT-3 available until the release of GPT-3.5-TURBO in March 2023. We do not experiment with models tuned for chat using reinforcement learning, nor with smaller but more efficient LLMs such as T5. Because the goal of this work is to generate translations, rather than to measure the acceptability of pre-existing translations, we sample strings from the model rather than measuring their probability (Hu and Levy, 2023); we use the standard text completion API with temperature .7 and 128 maximum tokens.

To translate an Inuktitut word into an English phrase, we look up every morph in the canonical segmentation and obtain their dictionary definitions. We then produce one or more prompts for the GPT-3 system. Figure 1 shows a sample prompt for the **Definition+example** method; examples of the other prompts are given in Appendix A.

Concatenate: As a trivial baseline, we simply concatenate the morpheme definitions in order, without using the LLM. This provides a point of comparison for evaluating the improvements due to text-to-text generation.

-Dictionary: GPT-3 has the capacity to translate many languages in a few-shot setting, and might have been exposed to definitions of Inuktitut words from the same web resources we are using. We use a prompt modeled on French-to-English few-shot translation (without dictionary definitions of morphemes, but with definitions of our few-shot words) to test whether this setting also works for Inuktitut.

Our next two methods evaluate the usefulness of specific parts of the dictionary entries. In the **definition only** condition, we provide textual definitions of each morpheme; if a morpheme matches multiple dictionary entries, we concatenate them with ";

or, " as the separator. In the **definition+example** condition, we also provide the English translation of an example word in which the morpheme is used (if the dictionary contains one).

We evaluate two more sophisticated methods for dealing with morphemes with multiple possible meanings. First, we ask GPT-3 textually to list all possible meanings for the word, rather than producing only a single one; we call this setting **multianswer**. Second, we enumerate all combinations of morpheme meanings which could make up the word, and create a separate prompt for each one. (This method requires much more computing time than simply prompting for more than one answer.) We call this setting **multiprompt**. In each case, the separate answers are aggregated in two different ways—average and oracle performance—as described in Section 3.

We consider two methods for injecting grammatical information into GPT-3's processing. First, we preface the morpheme decomposition with a short hand-written **grammar description**. Our grammatical description is intended to focus GPT-3 on some common issues we noticed in development. It explains that Inuktitut words begin with a root morpheme which usually determines the syntactic type of the word. Verbs must be translated as English sentences whose subject and object are given by agreement markers at the end of the word, while nouns must be translated as noun phrases or prepositional phrases. We also explain that intermediate morphemes can change the part of speech and contribute other elements to the meaning.

We also experiment with a **chain-of-thought** method in which the system is instructed to explicitly identify the syntactic category of the root, the category of the target translation, the subject and object (if any) from agreement morphology, and the function of any intermediate morphemes before translating.

We evaluate all the prompts, except the chain-of-thought and multiprompt methods, in both zero-shot and few-shot settings. The chain-of-thought method is used only in few-shot mode, since this allows us to model what the intermediate reasoning steps should look like. The multiprompt method is used only in zero-shot mode as, since a prompt is generated for each combination of definitions, it is extremely expensive to run with longer prompts. Our few-shot prompts are always filled in with a pre-selected list of the same five words, with definitions and grammatical decompositions from an Inuktitut pedagogy site, `tusaalanga.ca`. Three of these are translated as sentences, one as a noun phrase and one as a locative prepositional phrase. Two of the sentences have intransitive subject agreement markers and one has transitive subject-object agreement. We fill out the possible answers in the multi-answer condition and the chain of thought reasoning steps manually based on Tusaalanga.

4.1 Results

Table 4 shows the results. Overall, metric scores are low. Confidence intervals[3] are also very wide given the small size of the test set.

Existing NMT systems for Inuktitut can score around 30%, although these use more data and are not tested on exclusively multimorphemic words. BLEU scores in the 30s reflect generally intelligible though sometimes errorful translations; scores in the 20s are considered potentially useful under some circumstances, while not entirely accurate nor fluent. BLEURT scores, meanwhile, range between 0 and 1. Garcia et al. (2023) provides BLEURT scores for a variety of few-shot translation models. Scores for high-resourced German and Chinese are roughly 0.63-0.77; for less-resourced Icelandic they are 0.60-0.76.

Despite these caveats, some trends in the scores are evident. First, the scores of the **-Dictionary** condition compared to the rest show that GPT-3 has no useful prior knowledge of Inuktitut. The trivial **Concatenative** system scores higher, producing output which has some resemblance to the references, but is outperformed by the LLM systems, since it cannot rearrange content from the definitions into fluent translations.

Examining the non-trivial systems, we see that it is helpful to gather examples of morphemes in use, as well as definitions, from dictionary entries; these improve scores in both zero-shot and few-shot settings. Comparing the **multianswer** and **multiprompt** settings, we find that it is not very helpful to create multiple prompts to deal with polysemous morphemes; GPT-3 can handle polysemy naturally if asked to create multiple definitions. Finally, we find that the grammar lesson is unhelpful;

[3] Because BLEU scores represent global precision across the entire test set, we do not compute confidence intervals. We compute BLEURT confidence intervals using the SCIPY bootstrap method applied to the scores of each individual sentence.

Sys	BLEU	BLEURT (95% conf.)
Concat	6.19	0.43 (0.41 - 0.44)
-Dict	0.44	0.13 (0.12 - 0.14)
Def	9.63	0.48 (0.45 - 0.51)
Def+ex	13.29	0.51 (0.48 - 0.54)
Multians-avg	11.31	0.46 (0.44 - 0.48)
Multians-orac	19.83	0.62 (0.59 - 0.64)
Multipr-avg	17.42	0.50 (0.47 - 0.53)
Multipr-orac	23.30	0.59 (0.56 - 0.62)
Grammar	12.46	0.48 (0.45 - 0.51)
Few-shot		
Def	13.48	0.49 (0.46 - 0.52)
Def+ex	16.65	0.52 (0.48 - 0.55)
Multians-avg	18.47	0.51 (0.48 - 0.54)
Multians-orac	20.18	0.54 (0.51 - 0.57)
Grammar	17.30	0.53 (0.49 - 0.56)
Chain	13.91	0.43 (0.40 - 0.47)

Table 4: Metric scores for single-word translation. BLEURT scores are followed by bootstrapped 95% confidence intervals.

it is comparable to definitions and examples only in both zero-shot and few-shot settings. The **chain-of-thought** method, meanwhile, is actively unhelpful.

To gain more insight into the results, we show some translations of selected words in Table 5; the table contains two long verbs, one noun and one locative. The **-Dictionary** translations bear no resemblance to the references; although GPT-3 produces plausible and confident-seeming output, the meanings are completely confabulated. We do not observe wholesale hallucination in the definitions using the dictionary, although some grammatical features can be added incorrectly. For instance, the **definition only** system interprets the first word as a question despite the absence of any interrogative marker.

The **chain-of-thought** system has a tendency to lose information due to incomplete deductions. In example #2 (*tuktuliaqsimajut*), it identifies *tuktu* "caribou" correctly as a noun, but then fails to identify the verbalizing morpheme *liaq* "hunt"; because of this, it then states that there is no subject or object because the translation must be nominal and fails to translate the subject marker *jut* "they (3+)."

The other prompting strategies yield translations which are more similar to one another. In many cases, deviations from the reference reflect legitimate information from the definitions: in Example #1, *savik* can mean both "metal; steel" or "snow-knife". In #3, *sana* is defined as "to work at something; to fabricate, make something". In many cases, the translations obtained seem potentially useful for practical purposes.

kisarvik is an Inuktitut word which means "a place to anchor a boat". It is made up of the following parts:
kisaq: (currently unknown)
vik: place where the action of the verb takes place, as in "hospital; nursing station"; or, finality: 'for good'; 'forever', as in "He/she is leaving for good."; or, marks something that is immense or impressive in size, as in "ocean"
kisaq means:

Figure 2: An example prompt for definition elicitation.

On the other hand, morphemes with multiple meanings can lead to mistranslations. Uqausiit defines *unga* as "(root) to long or yearn for a person or a living thing"; "(root) the far side, the beyond of something"; "(locative) to (a place/location)". In the context of #4 *uvunga*, the locative meaning is applicable, since *unga* appears as a suffix, but this is apparently not sufficiently explained by our prompts. In addition, some of the systems appear to conflate information from multiple definitions of the morpheme.

5 Definition creation

We experiment with a single prompt for eliciting definitions. This prompt (Figure 2) provides the definition of all but a single morpheme, and the translation of the phrase as a whole, then asks for the definition of the missing item. GPT-3 is somewhat less likely to restrict itself to the prompted format in this case. We cut off elicited definitions at the first newline. For 20 of the roots, and 3 functional morphemes, GPT-3 repeats the prompt phrase "currently unknown."

Table 6 shows the results, which are much poorer than those for translation. This is partly due to the wide stylistic range of the definitions— reference definitions may contain more or fewer alternative synonyms, so that it is difficult to predict the correct length. However, the results of the task are also genuinely less reliable.

Inspection of the definitions (Table 7) echoes the numerical results, revealing some potential but also problematic tendencies. The system produces a correct definition of *aullaq* "leave" in Example #1, and arguably of *ijaq* "be cold" in #2. Such definitions could provide a starting point for a native speaker to rapidly expand a dictionary with new entries.

Inuktitut	#1 saviggirunnaqtutit	#2 tuktuliaqsimajut	#3 sanaji	#4 uvunga
Reference	You can bring your knife.	They are out caribou hunting.	a worker	to this spot here
-Dict	We are thankful (0.17)	We are learning (0.11)	I understand (0.04)	Peace (0.08)
Def only	are you able to bring a snow-knife? (0.52)	they have gone caribou hunting (0.74)	worker (0.36)	longing for (0.06)
Def+ex	you are able to bring a snow-knife (0.48)	they have gone caribou hunting (0.74)	a worker; a maker (0.61)	longing for (0.06)
Multianswer	You are able to bring metal along (0.39)	They (3+) have gone caribou hunting (0.62) They (3+) are away caribou hunting (0.62)	worker (0.36) maker (0.05) fabricator (0.03)	longing for something near here (0.55) yearning for something far away (0.14)
Grammar	you are able to bring a snow-knife (0.48)	they have gone hunting caribou (0.74)	a worker; someone who works (0.65)	longing/yearning for a place (0.39)
Chain	you can make it (0.23)	hunted caribou (0.31)	worker (0.36)	long for here (0.38)

Table 5: Examples of translations by -Dict and few-shot systems. Parenthesized numbers are BLEURT scores.

Data	BLEU	BLEURT (95% conf.)
Roots	2.88	0.33 (0.30 - 0.36)
Func.	2.97	0.27 (0.24 - 0.29)

Table 6: Metric scores for definition induction. BLEURT scores are followed by bootstrapped 95% confidence intervals.

On the other hand, the definitions of *qingaq* “nose” in #2 and *kisaq* “to anchor” in #3 are motivated by the provided examples, but too specific. *kisarvik* (#3), for instance, is made up of the target morpheme *kisaq* and *vik*, which creates a place nominal from a verb. The system therefore should infer that the target morpheme is a verb and does not contribute the meaning element “place”. Instead, it proposes the nominal meaning “an anchorage or place to tie up a boat.”

The system also hallucinates some entirely unmotivated definitions, such as “bay, inlet or cove” for *vik*. This definition is likely taken from the system’s knowledge of a different language such as Norwegian or Swedish, in which *vik* has this meaning. The extra work of filtering out these misleading definitions detracts from any benefit of using the system for dictionary expansion, since more time may be wasted than gained. These proposed definitions are also not suitable for presentation to a language learner, who may be misled by their seeming plausibility.

6 Conclusion

Our results clearly establish that GPT-3’s capacity for monolingual text understanding and generation are sophisticated enough to allow it to translate some complex words with the aid of a dictionary alone. Investigation of various prompt tuning techniques leads to a few tentative conclusions. Usage examples from the dictionary are helpful in combination with definitions. So are few-shot examples (even when only 5 are available). We investigated asking GPT-3 to generate multiple translations, and found some evidence that it might be helpful, particularly when the user already has some idea of how the word might be disambiguated based on the context in which it was encountered. Experiments on better-resourced agglutinative languages, such as Finnish and Hungarian, would be useful in better differentiating the performance of different prompting techniques.

On the other hand, we were not able to show much benefit from adding information on grammar to the prompts. There are two possible reasons: Perhaps we engineered the prompts poorly and could have obtained better results with more careful decisions about what information to include. Or perhaps the GPT-3 system is not yet sophisticated enough to relate relatively abstract grammar information to the concrete task of translation.

We believe both of these explanations are valid to some extent. Evidence that our prompt engineering is at fault comes from examining the failures of the chain-of-thought system in more detail (Appendix B). As stated above, the system sometimes misses nominalizers and verbalizers; although the prompt states that these exist, we could have included more detailed information on how to recognize them.

On the other hand, the system also sometimes creates complex chains of thought in which it loses track of content morphemes. This is why it predicts the meaning “you can make it” for #1 in Table 5 rather than “you can bring your knife”, forgetting the initial element *savik*. Similar problems apply to some words expressing emotion— *tautugusunn-*

Example	#1 aullarli "Let him/her depart!"	#2 qingaijaqtunga "I have a cold nose."	#3 kisarvik "a place to anchor a boat"	#4 kanunga "to that spot down there"
Content (root) morphemes				
Inuktitut	aullaq	qingaq	kisaq	kan
Reference	to leave; to depart; to go away	nose	to anchor something	a place/spot down here/there
System	to go away; to leave; to depart; to move away; to go out. (0.74)	to be cold, shivering, chilly or having a cold nose (0.31)	an anchorage or place to tie up a boat (0.37)	that (0.02)
Functional morphemes				
Inuktitut	li	ijaq	vik	unga
Reference	let him/her/it...! (command)	(1) to remove; to have something removed (2) to experience coldness of body parts	place where the action of the verb takes place	to (a place/location)
System	him/her (0.28)	to be cold (0.35)	bay, inlet, or cove (0.26)	to go, as in "to go down there" (0.33)

Table 7: Examples of definitions induced by a one-shot system morphemes, with BLEURT scores.

gittara "I don't feel like watching it" is translated as "I love him/her/it" because the system expands *gusuk* "feel an emotion" into "love", replacing the legitimate main verb "watch".

While better prompting could potentially remedy some of these issues, we believe more sophisticated instruction-following models may perform even better given the same resources. Moreover, models with very long context windows might be able to read in large sections of pedagogical material from sites like Tusaalanga directly, reducing the necessity to create abbreviated grammar lessons specifically for use in prompting LLMs. We are also interested to see to what extent more sophisticated LLMs can improve at suggesting dictionary definitions, especially to the extent that hallucinations can be controlled. On the other hand, it would be potentially interesting to see whether these results can be replicated with smaller, more cost and computation-effective LLMs such as T5.

While this work emphasizes how much is possible without an NMT system, we also believe that it can contribute to NMT development in the very low-resource case. Curriculum learning for translation (e.g. Platanios et al., 2019) uses translations of shorter or simpler constructions earlier in pretraining, but suitable "easy" instances may be rare in corpus data, or their distribution may be skewed towards formulaic language. Few-shot dictionary-based translation could be used to bootstrap towards a larger NMT system by providing candidate definitions for single words from the corpus.

Although the scope of the present work is limited to an exploratory demonstration, we are eager to see the many ways in which it can be expanded upon. One particular direction is to explore the extent to which an LLM can exploit linguistic context to disambiguate between various potential translations, hopefully leading to a narrowing of the gap between average and oracle performances. More closely integrating segmentation into the prompting system, either by having the LLM produce its own segmentations or rank multiple segmentations based on the plausibility of their meanings, would reduce dependence on an accurate canonical segmentation system.

We are (to our knowledge) the first to evaluate dictionary-based translation in the absence of a base NMT system, and the first to deploy it on a polysynthetic language. While our results are not yet competitive with fully trained translation, we believe our results represent good news for communities in which limited resources must be distributed among efforts to develop community-facing resources or parallel corpora. A community that focuses its effort on developing dictionaries for human learners can nonetheless enjoy some of the benefits of MT without developing a conventional NMT system, helping to bring language revitalization and language technologies closer together.

Limitations

The results of this work may be limited in reliability and replicability due to some hard-to-avoid aspects of the low-resource setting.

Our numerical results have low statistical power, as illustrated by the wide BLEURT confidence intervals in Tables 4 and 6. Without a large test set, most differences are not statistically significant at the accepted level. They should be treated as trends which can motivate further investigation rather than solid conclusions. The significant findings are that the no dictionary system is worse than the concatenative baseline, which is in turn worse than the LLM systems; multianswer and multiprompt oracles surpass the definition-only system in zero and few-shot settings.

Human evaluations would also improve the reliability of our evaluations, which are currently entirely automated. However, we do not have access to Inuktitut native speakers. Human evaluations could also be used to improve the automated metrics by fine-tuning BLEURT.

Reproducibility of our experiments is limited by potential changes to the OpenAI models we use; OpenAI might withdraw or update them at any time. We estimate that the project incurred total costs under $70 in payments for the GPT-3 API. The multiprompt experiment (which generates a prompt for each combination of morpheme definitions, and did not meaningfully improve over asking GPT-3 to provide multiple answers) was responsible for much of this cost. Few-shot prompts are also more expensive than zero-shot due to their length. We believe that only a few dollars would be necessary to reproduce the most successful system here and run it on hundreds of examples.

Finally, our method assumes access to a canonical segmentation system, which potentially limits its applicability to very low-resourced languages where such a system may be unavailable. By filtering out incorrectly segmented examples, we do not assess the potential impact of segmentation errors on translation.

Ethics Statement

We reached out to the Inuit Uqausinginnik Taiguusiliuqtiit with regard to their stance on extracting datasets from Uqausiit but have not received any reply. We therefore do not plan to make the datasets available to the community for download.

If a system for word translation based on this paper were deployed, it should be in the context of clear labeling. It would be important to indicate the morpheme analysis and morpheme definitions for the word being analyzed, and clearly separate the automatically generated proposed translation, which should be designated as the product of a system which lacks native-speaker expertise. Because most systems in this task did not hallucinate definitions, we believe that a clearly labeled system of this type might do more good than harm in the context of a revitalization effort.

If a system for definition induction based on this paper were deployed, it would be extremely important that only native speakers were allowed to use computer-authored definitions as sources for dictionary entries, and that they be told clearly that the system was provided only as a labor-saving device, rather than as a source of native-like expertise. Scots Wikipedia is one widely cited case where a naive user added a large amount of misleading data to an online resource under the impression that they were being helpful (Brooks and Hern, 2020). Our definition induction system has the potential for this kind of misuse.

Acknowledgements

We thank the Clippers computational discussion group for comments on a preliminary version of this work. We thank four anonymous reviewers for their suggestions, including feedback which helped us to correct a serious problem with our evaluation procedures.

References

Loïc Barrault, Magdalena Biesialska, Ondřej Bojar, Marta R. Costa-jussà, Christian Federmann, Yvette Graham, Roman Grundkiewicz, Barry Haddow, Matthias Huck, Eric Joanis, Tom Kocmi, Philipp Koehn, Chi-kiu Lo, Nikola Ljubešić, Christof Monz, Makoto Morishita, Masaaki Nagata, Toshiaki Nakazawa, Santanu Pal, Matt Post, and Marcos Zampieri. 2020. Findings of the 2020 conference on machine translation (WMT20). In Proceedings of the Fifth Conference on Machine Translation, pages 1–55, Online. Association for Computational Linguistics.

Libby Brooks and Alex Hern. 2020. Shock an aw: Us teenager wrote huge slice of scots wikipedia. nineteen-year-old says he is "devastated" after being accused of cultural vandalism. The Guardian, 26.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. 2020. Language models are few-shot learners. Advances in neural information processing systems, 33:1877–1901.

Georgiana Dinu, Prashant Mathur, Marcello Federico, and Yaser Al-Onaizan. 2019. Training neural machine translation to apply terminology constraints. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages

3063–3068, Florence, Italy. Association for Computational Linguistics.

Zi-Yi Dou and Graham Neubig. 2021. Word alignment by fine-tuning embeddings on parallel corpora. In Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, pages 2112–2128, Online. Association for Computational Linguistics.

Benoit Farley. 2012. Uqailaut Inuktitut morphological analyzer.

Xavier Garcia, Yamini Bansal, Colin Cherry, George Foster, Maxim Krikun, Fangxiaoyu Feng, Melvin Johnson, and Orhan Firat. 2023. The unreasonable effectiveness of few-shot learning for machine translation.

Nikesh Garera and David Yarowsky. 2008. Translating compounds by learning component gloss translation models via multiple languages. In Proceedings of the Third International Joint Conference on Natural Language Processing: Volume-I.

Marjan Ghazvininejad, Hila Gonen, and Luke Zettlemoyer. 2023. Dictionary-based phrase-level prompting of large language models for machine translation.

Stig-Arne Grönroos, Katri Hiovain, Peter Smit, Ilona Erika Rauhala, Päivi Kristiina Jokinen, Mikko Kurimo, and Sami Petteri Virpioja. 2016. Low-resource active learning of morphological segmentation. Northern European Journal of Language Technology.

Eva Hasler, Felix Stahlberg, Marcus Tomalin, Adrià de Gispert, and Bill Byrne. 2017. A comparison of neural models for word ordering. In Proceedings of the 10th International Conference on Natural Language Generation, pages 208–212, Santiago de Compostela, Spain. Association for Computational Linguistics.

Jennifer Hu and Roger Levy. 2023. Prompt-based methods may underestimate large language models' linguistic generalizations. Lingbuzz preprint.

Eric Joanis, Rebecca Knowles, Roland Kuhn, Samuel Larkin, Patrick Littell, Chi-kiu Lo, Darlene Stewart, and Jeffrey Micher. 2020. The Nunavut Hansard Inuktitut–English parallel corpus 3.0 with preliminary machine translation results. In Proceedings of the Twelfth Language Resources and Evaluation Conference, pages 2562–2572, Marseille, France. European Language Resources Association.

Chen Liang, Haoming Jiang, Xiaodong Liu, Pengcheng He, Weizhu Chen, Jianfeng Gao, and Tuo Zhao. 2021. Token-wise curriculum learning for neural machine translation. In Findings of the Association for Computational Linguistics: EMNLP 2021, pages 3658–3670, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Zoey Liu, Robert Jimerson, and Emily Prud'hommeaux. 2021. Morphological segmentation for Seneca. In Proceedings of the First Workshop on Natural Language Processing for Indigenous Languages of the Americas, pages 90–101, Online. Association for Computational Linguistics.

Thang Luong, Ilya Sutskever, Quoc Le, Oriol Vinyals, and Wojciech Zaremba. 2015. Addressing the rare word problem in neural machine translation. In Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pages 11–19, Beijing, China. Association for Computational Linguistics.

Manuel Mager, Özlem Çetinoğlu, and Katharina Kann. 2020. Tackling the low-resource challenge for canonical segmentation. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 5237–5250, Online. Association for Computational Linguistics.

Manuel Mager, Ximena Gutierrez-Vasques, Gerardo Sierra, and Ivan Meza-Ruiz. 2018a. Challenges of language technologies for the indigenous languages of the Americas. In Proceedings of the 27th International Conference on Computational Linguistics, pages 55–69, Santa Fe, New Mexico, USA. Association for Computational Linguistics.

Manuel Mager, Elisabeth Mager, Alfonso Medina-Urrea, Ivan Vladimir Meza Ruiz, and Katharina Kann. 2018b. Lost in translation: Analysis of information loss during machine translation between polysynthetic and fusional languages. In Proceedings of the Workshop on Computational Modeling of Polysynthetic Languages, pages 73–83, Santa Fe, New Mexico, USA. Association for Computational Linguistics.

Mick Mallon. 2000. Inuktitut linguistics for technocrats.

Jeffrey Micher. 2017. Improving coverage of an Inuktitut morphological analyzer using a segmental recurrent neural network. In Proceedings of the 2nd Workshop on the Use of Computational Methods in the Study of Endangered Languages, pages 101–106, Honolulu. Association for Computational Linguistics.

Simon Mille, Anya Belz, Bernd Bohnet, Thiago Castro Ferreira, Yvette Graham, and Leo Wanner. 2020. The third multilingual surface realisation shared task (SR'20): Overview and evaluation results. In Proceedings of the Third Workshop on Multilingual Surface Realisation, pages 1–20, Barcelona, Spain (Online). Association for Computational Linguistics.

Sarah Moeller and Mans Hulden. 2021. Integrating automated segmentation and glossing into documentary and descriptive linguistics. In Proceedings of the 4th Workshop on the Use of Computational Methods in the Study of Endangered Languages Volume 1 (Papers), pages 86–95, Online. Association for Computational Linguistics.

Tan Ngoc Le and Fatiha Sadat. 2020. Revitalization of indigenous languages through pre-processing and neural machine translation: The case of Inuktitut. In Proceedings of the 28th International Conference on Computational Linguistics, pages 4661–4666, Barcelona, Spain (Online). International Committee on Computational Linguistics.

Jan Niehues. 2021. Continuous learning in neural machine translation using bilingual dictionaries. In Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume, pages 830–840, Online. Association for Computational Linguistics.

Hyunji Hayley Park, Katherine J. Zhang, Coleman Haley, Kenneth Steimel, Han Liu, and Lane Schwartz. 2021. Morphology matters: A multilingual language modeling analysis. Transactions of the Association for Computational Linguistics, 9:261–276.

Ngoc-Quan Pham, Jan Niehues, and Alexander Waibel. 2018. Towards one-shot learning for rare-word translation with external experts. In Proceedings of the 2nd Workshop on Neural Machine Translation and Generation, pages 100–109, Melbourne, Australia. Association for Computational Linguistics.

Emmanouil Antonios Platanios, Otilia Stretcu, Graham Neubig, Barnabas Poczos, and Tom Mitchell. 2019. Competence-based curriculum learning for neural machine translation. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 1162–1172, Minneapolis, Minnesota. Association for Computational Linguistics.

Matt Post. 2018. A call for clarity in reporting BLEU scores. In Proceedings of the Third Conference on Machine Translation: Research Papers, pages 186–191, Brussels, Belgium. Association for Computational Linguistics.

Christian Roest, Lukas Edman, Gosse Minnema, Kevin Kelly, Jennifer Spenader, and Antonio Toral. 2020. Machine translation for English–Inuktitut with segmentation, data acquisition and pre-training. In Proceedings of the Fifth Conference on Machine Translation, pages 274–281, Online. Association for Computational Linguistics.

Thibault Sellam, Dipanjan Das, and Ankur Parikh. 2020. BLEURT: Learning robust metrics for text generation. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 7881–7892, Online. Association for Computational Linguistics.

Zewei Sun, Qingnan Jiang, Shujian Huang, Jun Cao, Shanbo Cheng, and Mingxuan Wang. 2022. Zero-shot domain adaptation for neural machine translation with retrieved phrase-level prompts.

Adam Wiemerslage, Miikka Silfverberg, Changbing Yang, Arya McCarthy, Garrett Nicolai, Eliana Colunga, and Katharina Kann. 2022. Morphological processing of low-resource languages: Where we are and what's next. In Findings of the Association for Computational Linguistics: ACL 2022, pages 988–1007, Dublin, Ireland. Association for Computational Linguistics.

Tong Zhang, Long Zhang, Wei Ye, Bo Li, Jinan Sun, Xiaoyu Zhu, Wen Zhao, and Shikun Zhang. 2021. Point, disambiguate and copy: Incorporating bilingual dictionaries for neural machine translation. In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), pages 3970–3979, Online. Association for Computational Linguistics.

Xuan Zhang and Kevin Duh. 2021. Approaching sign language gloss translation as a low-resource machine translation task. In Proceedings of the 1st International Workshop on Automatic Translation for Signed and Spoken Languages (AT4SSL), pages 60–70, Virtual. Association for Machine Translation in the Americas.

Yuhao Zhang, Ziyang Wang, Runzhe Cao, Binghao Wei, Weiqiao Shan, Shuhan Zhou, Abudurexiti Reheman, Tao Zhou, Xin Zeng, Laohu Wang, Yongyu Mu, Jingnan Zhang, Xiaoqian Liu, Xuanjun Zhou, Yinqiao Li, Bei Li, Tong Xiao, and Jingbo Zhu. 2020. The NiuTrans machine translation systems for WMT20. In Proceedings of the Fifth Conference on Machine Translation, pages 338–345, Online. Association for Computational Linguistics.

A Example prompts

We show the complete zero-shot prompt of each type for Example #1 of Table 5, *saviggirunnaqtutit* "You can bring your knife." The outputs are shown in the table.

No dictionary:

Translate Inuktitut to English:
saviggirunnaqtutit =>

Definition only:

The Inuktitut word saviggirunnaqtutit is made up of the following parts:
savik: (1) metal; steel; iron (2) snow-knife; or, to be fitted with a metal point (harpoon; spear)
ggiq: to bring someone or something along

runnaq: to be able to perform a certain action

tutit: you

In English, saviggirunnaqtutit means roughly:

Definition+example:

The Inuktitut word saviggirunnaqtutit is made up of the following parts:

savik: (1) metal; steel; iron (2) snow-knife; or, to be fitted with a metal point (harpoon; spear)

ggiq: to bring someone or something along, as in "Don't bring your gun along."

runnaq: to be able to perform a certain action, as in "Could you find that out if he/she arrived"

tutit: you, as in "you sleep"

In English, saviggirunnaqtutit means roughly:

Multianswer:

The Inuktitut word saviggirunnaqtutit is made up of the following parts:

savik: (1) metal; steel; iron (2) snow-knife; or, to be fitted with a metal point (harpoon; spear)

ggiq: to bring someone or something along, as in "Don't bring your gun along."

runnaq: to be able to perform a certain action, as in "Could you find that out if he/she arrived"

tutit: you, as in "you sleep"

Give all possible translations of saviggirunnaqtutit:

Grammar lesson:

An Inuktitut word is made up of a root and some optional modifiers; for verbs, this will be followed by a verb ending which acts as an agreement marker.

If the root is a verb, the whole word will usually be translated as a sentence. If the root is a noun, the whole word will be translated as a noun phrase or a prepositional phrase. If the word ends with a locative modifier (like "in" or "on"), translate it as a prepositional phrase.

Some words contain a nominalizer which turns a verb into a noun, like "someone who does the action" or "location where the action takes place". These should be translated as noun phrases even though the root is a verb.

If the translation is a sentence, its subject (and object if there is one) will be given by a verb ending.

Other modifiers within the sentence can introduce auxiliary verbs, adverbs or discourse particles.

The material above is repeated only once in the few-shot setting; the material below is copied for each few-shot word.

The Inuktitut word saviggirunnaqtutit is made up of the following parts:

savik: (1) metal; steel; iron (2) snow-knife; or, to be fitted with a metal point (harpoon; spear)

ggiq: to bring someone or something along, as in "Don't bring your gun along."

runnaq: to be able to perform a certain action, as in "Could you find that out if he/she arrived"

tutit: you, as in "you sleep"

In English, saviggirunnaqtutit means roughly:

Chain-of-thought:

Since we use chain-of-thought prompting only in few-shot mode, we show a few-shot prompt.

An Inuktitut word is made up of a root and some optional modifiers; for verbs, this will be followed by a verb ending which acts as an agreement marker.

To translate an Inuktitut word, first, identify the part of speech of the root.

If the root is a verb, the whole word will usually be translated as a sentence.

If the root is a noun, the whole word will be translated as a noun phrase or a prepositional phrase. If the word ends with a locative modifier (like "in" or "on"), translate it as a prepositional phrase.

Some words contain a nominalizer which turns a verb into a noun, like "someone who does the action" or "location where the action takes place". These should be translated as noun phrases even though the root is a verb.

State the syntactic category of the translation.

If the translation is a sentence, its subject (and object if there is one) will be given by a verb ending. Using this ending, state the subject and object.

Other modifiers within the sentence can introduce auxiliary verbs, adverbs or discourse particles. State the meaning of each modifier.

Finally, translate the word into English.

The Inuktitut word aquttunnaqtuq is made up of the following parts:

aqut: to steer or drive a vehicle or boat, as in "Who is going to drive?"

junnaq: to be able to perform a certain action, as in "He/she can hear."

juq: he/she/it, as in "he/she/it sees"

The root aqut is a verb.

The translation will be a sentence.

The subject is he/she/it and there is no object.

junnaq means 'can', creating the meaning "can drive"

The translation is => he/she/it can drive

The Inuktitut word quviasuppit is made up of the following parts:

quviak: to be happy, joyful, as in "they were happy while they did something"

suk: added to verb roots that normally are transitive (double) to make them intransitive (single), as in "He/she is afraid."

vit: the...of your..., as in "the window of your (1) house"; or, Are you...?; Do you...?, as in "Are you eating?"

The root quviak is a verb.

The translation will be a sentence.

The subject is you and there is no object.

suk makes the verb intransitive, creating the meaning "be happy"

The translation is => are you happy?

The Inuktitut word maligaliurvik is made up of the following parts:

malik: (1) to follow someone or something) (2) to obey someone or something, as in "Kiviuq followed a person into the tent."

gaq: changes a verb to a noun with a passive meaning: 'something that one...', as in "a drink of something; a soft drink"

liq: an action that is underway or starting; also marks a change from one state to another, as in "He/she is going out right now."

uq: marks a continuous, prolonged or repeated action, as in

"He pulled in the long rope."
vik: place where the action of the verb takes place, as in "hospital; nursing station"; or, finality: 'for good'; 'forever', as in "He/she is leaving for good."; or, marks something that is immense or impressive in size, as in "ocean"
The root malik is a verb.
The translation will be a noun phrase because of the nominalizer vik.
There is no subject or object because it is a noun phrase.
gaq makes the verb passive, creating the meaning "be obeyed"
liq and uq together indicate association with a habit or profession, creating the meaning "legislation"
vik indicates the place in which legislation happens, creating the meaning "legislative assembly"
The translation is => Legislative Assembly

The Inuktitut word allavvimmi is made up of the following parts:
allak: used to tell someone to do something quickly or for a particular reason, as in "Sleep a little! (before you do something else)"; or, to write or draw vik: place where the action of the verb takes place, as in "hospital; nursing station"; or, finality: 'for good'; 'forever', as in "He/she is leaving for good."; or, marks something that is immense or impressive in size, as in "ocean" mi: of his/her own...; or, of their (2)...; or, also; too, as in "I saw it, too."; or, of his/her own..., as in "the pole of his/her own tent"; or, in/on/at...(one of something), as in "He sleeps in the tent"
The root allak is a verb
The translation will be a prepositional phrase because of the nominalizer vik and the locative marker mi.
There is no subject or object because it is a noun phrase.
vik indicates the place in which writing happens, creating the meaning "office" The translation is "in/at the office"

The Inuktitut word saviggirunnaqtutit is made up of the following parts:
savik: (1) metal; steel; iron (2) snow-knife; or, to be fitted with a metal point (harpoon; spear)
ggiq: to bring someone or something along, as in "Don't bring your gun along."
runnaq: to be able to perform a certain action, as in "Could you find that out if he/she arrived"
tutit: you, as in "you sleep"
The root

B Chain-of-thought output

We show the full model output for the example words in Table 5; see the previous appendix for the prompt format.

saviggirunnaqtutit "You can bring your knife":
savik is a verb.
The translation will be a sentence.
The subject is you and there is no object.
ggiq and runnaq together indicate the ability to make something, creating the meaning "make"
The translation is => you can make it.

tuktuliaqsimajut "They are out caribou hunting":
tuktu is a noun.
The translation will be a noun phrase.
There is no subject or object because it is a noun phrase.
liaq indicates hunting and sima indicates the result of the hunting, creating the meaning "hunted caribou"
The translation is => hunted caribou.

sanaji "a worker"
sana is a verb.
The translation will be a noun phrase because of the nominalizer ji.
The subject is one who performs the action and there is no object.
ji indicates that the verb is turned into a noun, creating the meaning "worker"
The translation is => worker

uvunga "to this spot here": Is the font size change here significant?
unga is a verb.
The translation will be a prepositional phrase because of the locative marker uv.
There is no subject or object because it is a noun phrase.
unga indicates a sense of longing, creating the meaning "long for"
The translation is => long for here.