Learning isolated polysemous words: identifying the intended meaning of language learners in informal ubiquitous language learning environments

Victoria Abou-Khalil[1*] ⓘD, Samar Helou[1], Brendan Flanagan[2], Mei-Rong Alice Chen[2] and Hiroaki Ogata[2]

*Correspondence:
v.aboukhalil@gmail.com
[1]Kyoto University, Department of
Social Informatics, Graduate School
of Informatics, Kyoto, Japan
Full list of author information is
available at the end of the article

Abstract

A growing number of language learners use ubiquitous language learning applications to learn anytime and anywhere. Learners translate and learn isolated words inspired by their activities and surroundings. However, isolated words may have several meanings that change depending on the context. Since learners don't have the opportunity to indicate the meaning they are looking for in an online learning environment, they risk learning translations that do not correspond to their intended meaning. Identifying the intended meaning of the learner is needed to provide them with an appropriate translation. However, isolated words are difficult to disambiguate due to a lack of text around them. To this end, informal ubiquitous learning environments can offer another type of context, one that is formed by the users' past learning logs. In this work, we propose using the learners' past vocabulary to disambiguate their intended meaning when they look up isolated words. Accordingly, we propose and evaluate three methods. The first method considers that the intended meaning of the learner is the one that is the most semantically similar to the learner's past vocabulary. The second method builds on the first method but gives more weight to the vocabulary that the learner logged shortly before the target word. The third method addresses situations where the semantic similarities between the different meanings of the word and the past vocabulary have similar values. In those cases, the method considers that the intended meaning of the learner is the most common meaning in the target language. The three methods were evaluated using 148 logs of SCROLL, a ubiquitous informal language learning environment. The success rates of the three methods were 72.180%, 75.630%, and 83.050% respectively. This work shows that the past activity of language learners in informal ubiquitous language learning environments could be used to identify their intended meaning when learning a new word.

Keywords: Language learning, Polysemy, Homographs, Vocabuary, Word sense disambiguation, Computer assisted language learning, Ubiquitous learning systems

Introduction

A growing number of language learners are using informal language learning applications to learn new vocabulary anytime and anywhere. Dictionaries and online translators are learners' preferred language learning tools (Demouy et al., 2016) with Google Translate being the most used translation tool (Ducar & Schocket, 2018). Mobile translators or

informal vocabulary learning systems allow language learners to input a word they wish to learn, get the translation, and save it as a log for future review sessions. Learners encounter issues with translation when they encounter words that can have several meanings depending on the context. These are known as polysemous words and homographs. A polysemous word is a word that has different meanings that derive from a common origin; a homograph is a word that has different meanings with unrelated origins. Polysemous words and homographs constitute a known problem for language learners.

Online translators usually provide users with the translation of only one of the meanings of the target word. Learners do not have the opportunity to indicate the meaning of the word they are looking up when learning vocabulary in an informal language learning environment. This can cause learners to learn words out of their context of use, eventually leading to miscommunication. In fact, language learners have difficulties knowing the correct translation for their intended meaning (Boulton & De Cock, 2017). Even when using a dictionary, learners tend to look at the top entry, and will rarely refer to other entries (Jin & Deifell, 2013). It is important to provide the learners the translation they are looking for better comprehension of vocabulary. To provide the learners with the translation they want, we need to first identify their intended meaning when they translate isolated polysemous words.

Classical Word Sense Disambiguation (WSD) techniques are unable to disambiguate isolated words as they require a text or sentences surrounding the target word to give a context for the disambiguation. In this respect, informal ubiquitous language learning applications could provide us with a new kind of context to identify a learner's intended meaning. In fact, language learners tend to learn words inspired by their activities or surroundings (Sharples et al., 2005). Moreover, informal ubiquitous vocabulary learning systems allow the gathering of the user's digital trace, which is formed by all the learning logs created by the user on the system. A learning log usually includes the word that has been learned, the location and time of learning, and a photo or video associated with the word. Learning logs have been used for contextual vocabulary learning purposes. Some typical use cases include teaching words to users in a specific location that other users learned in a similar location, or reminding users of words they learned previously when they return to a specific place (Ogata et al., 2011). This digital trace could also be used as a context to disambiguate isolated words and provide the appropriate translation to the learner. Even though the learners' digital trace is increasingly available, it has not yet been used as a context to identify the intended meaning of language learners.

When language learners look up an isolated word in a translation tool, or in an informal ubiquitous language learning environment that incorporates a translation tool, they don't have access to the surrounding text that would give the context needed to disambiguate the word. In this work, we propose to use the learners' past logs as a context to identify their intended meaning when they translate isolated homographs or polysemous words in an informal ubiquitous language learning environment. We propose three different methods that aim to identify the learner's intended meaning based on the semantic similarity between the learner's past vocabulary and the different meanings of the target word. The method is evaluated on users of the SCROLL system (System for Capturing and Reminding Of Learning Log) (Ogata et al., 2011). SCROLL is an informal language learning application in which users record the words they have learned in their daily lives. The proposed methods could enable the design of systems that provide language

learners with translations based on their intended meaning, essential to the improvement of their vocabulary learning and communication skills. Considering all of the above, the two objectives of this work can be stated as:

– to demonstrate that the logs generated in a ubiquitous language learning environment can be used to identify a learner's intended meaning when they look up or translate isolated polysemous words.
– to propose methods to identify a learner's intended meaning when they translate polysemous words in a ubiquitous language learning environment.

Background

Informal vocabulary learning using technologies

Vocabulary learning is a major component of language learning. Previous studies have shown that a lack of lexical knowledge is the biggest impediment to fluent spoken production of language and recommended that L2 learners learn as manyL2 vocabulary words as possible (Hilton, 2008). Moreover, the vocabulary size of the learners is reflected in their written production (Laufer & Nation, 1995). Another study found that the size of a learner's vocabulary is strongly associated with their reading and writing abilities (Stæhr, 2008). The use of technologies in language learning support lifelong and contextual vocabulary learning (Dimakopoulos & Magoulas, 2009) and is becoming more popular among language learners. A review on language learning applications showed that vocabulary instruction was the main focus of 84% of the reviewed language learning applications. Moreover, the majority of language learning applications focused on vocabulary in isolation, that is, individual words without context (Heil et al., 2016). More than 90% of surveyed language learners use their mobile phones informally for their learning activities. Moreover, other studies have shown that language learners' preferred language learning tools are dictionaries and online translation tools (Demouy et al. 2016). In fact, a survey of 900 language students was conducted in 2012 and showed that 71% of the students were using machine translation, with 89% using it as a dictionary (Mueller, 2013). The most widely used translation tool is Google Translate (Ducar & Schocket, 2018). Using these tools, learners choose the vocabulary they are about to learn. One important aspect of ubiquitous and self-directed language learning tools is that the context is constructed by the learners through interaction. Learners choose to learn words inspired by their surroundings, interests and goals (Sharples et al., 2005).

Polysemy and technology supported vocabulary learning

A recent study shows that second language learners better learn a word by associating it with the translation of that word in their first language, then by associating it to with their definition of the word in the second language (Joyce, 2018). In other terms, second language learners benefit from learning a new word with its equivalent translation in their first language. However, the meaning of a word changes from one context to another, as does its translation. In fact, about 40% of English words have more than one meaning listed in a dictionary (Nagy, 1995). This number includes polysemous words – words that have several related meanings–and homographs, i.e., words that have several unrelated meanings. In a dictionary, polysemous words have different meanings under the same entry, whereas homographs are listed in multiple entries. Polysemous words

and homographs can be problematic to self-directed language learners. In fact, learners have difficulties knowing the correct translation of a word or which meaning they should choose in order to use a word in a specific context (Boulton & De Cock, 2017). Of course, second language learners can always refer to a dictionary that provides detailed definitions as well as examples of usage. However, most learners prefer bilingual dictionaries (Atkins & Varantola, 1997; Boulton & De Cock, 2017) which usually lack rich lexico-grammatical information (Boulton & De Cock, 2017). Moreover, bilingual language dictionaries order the words by meaning frequency. Users tend to look at the top entry and will not refer to other entries unless the first one is obviously wrong. Google translate is the most used online tool amongst language learners (Jin & Deifell, 2013). Machine translation of single words provides the user with one translation that usually consists of the most common meaning. Using this kind of simplified bilingual dictionary or online translator can lead language learners to learn a different translation then the one appropriate to their intended context of use. Language learners would benefit from getting the appropriate translation in their intended context of use (Nagy, 1995). In a classroom environment, a learner has the opportunity to communicate their intended usage of a word with multiple meanings to the teacher, and subsequently learn the appropriate translation of the word. However, in an informal and distant learning environment, language learners cannot state the meaning of the word they are looking for. This is especially true when learners are looking up translations of single words. One obvious solution would be to provide the learner with a list of definitions and corresponding translations. However, if given such a list, learners may still find it difficult to identify the meaning they are looking for. Miller and Gildea suggest that definitions should be given in the context encountered by the learner during the moment of learning instead of in a list of different dictionary definitions (Miller & Gildea, 1987). Providing the learners with a translation that corresponds to their intended meaning requires adaptivity to individual users. Most of today's applications do not adapt to the needs of individual learners, and a more adaptive learning would provide a more personalized experience in terms of the content delivered during instruction (Heil et al., 2016).

However, dictionaries and ubiquitous vocabulary learning applications used today register the lifelong activity performed by language learners and can be considered a form of lifelogging. Lifelogging is defined as a digital record of what a learner has learned in their daily life using ubiquitous computing technologies. Lifelogging systems can be used to help understand the user and their intended meaning when they look up a word in the system. Lifelogging systems provide information on the user's environment and the information needs of the user. These systems give insight into what the user is experiencing and learning at any point in time and any moment prior to it. Lifelogging systems offer the potential to tailor information to the user in response to an information need (Gurrin et al., 2014).

Word sense disambiguation and vocabulary learning

Word Sense Disambiguation (WSD) is a well-established field and a common problem of natural language processing. The objective of WSD is the identification of the most proper (dictionary) definition for an ambiguous word in a given context. In computational linguistics, the context usually consists of the words and sentences that occur around the target word (De Brabanter & Kissine, 2009). WSD has been used for vocabulary learning

purposes. However, to our knowledge, all previous studies worked on disambiguating words that were encountered within a text; therefore their aim was to identify the meaning intended by the author of the text. WSD has been used to help learners with their reading comprehension and vocabulary learning by providing them context-specific definitions while they are reading a text (Eom et al., 2012; Azab et al., 2013). However, when language learners look up the definition or translation of isolated words, text surrounding polysemous words is not available to give context for disambiguation and learners might end up learning the translation that does not correspond to their intended context. In this work, we propose using learners' past logs to identify their intended meaning.

Methods

Data collection from an informal distance language learning environment

The informal language learning environment used in this study is the SCROLL system, (System for Capturing and Reminding Of Learning Log). SCROLL is a digital record of what language learners are learning in their daily lives. SCROLL's users create an account and log the vocabulary they wish to learn. SCROLL allows users to capture the contextual data when learning a new word – users can log the new word they learned, get its translation, save the time of insertion, an image, a video, and their current location (Ogata et al., 2011). SCROLL is free to use and currently has 1705 registered users and contains around 30380 logs (Ogata et al., 2018). 34.2% of the logs have a location associated to them. Most SCROLL users live in Japan and use SCROLL to informally learn either English or Japanese.

When a user wants to learn a new word, they create a new log, log_n. Unless the user is new to the system, their profile already contains their previous logs: $log_1, log_2, \ldots, log_{n-1}$. Each *log* is constituted by a word, an image, a location and a time of input.

$log = \{word, image, location, time\}$

Most of SCROLL's users do not upload an image when creating a new log. Only 18.6% of logs are created with images associated with them (Hasnine et al., 2018).

Even though a *word* can have more than one meaning, SCROLL displays only one translation to the user, using Microsoft Translator Text API. Figure 1 is a screenshot from SCROLL that shows a log inserted by a learner for the word *book*. The user attached an image and a location when creating the log. A Japanese and a Portuguese translation of the word *book* is automatically provided to the user, and the time of input is automatically saved.

Figure 2 shows the typical usage of SCROLL. The user saves a new word they want to learn, depending on their context . The translation is provided to the user in the target language. The time and location are saved automatically, and the user is able to attach an image to the log. The user can also have access to their previous logs as shown in Fig. 2.

General and situated vocabulary in the users' digital trace

To understand the patterns of vocabulary logging using an informal ubiquitous language learning environment, we selected 20 SCROLL users and analyzed their learning activity. We picked the users randomly from a larger user group that was selected based on the following criteria:

book
06/04/19

Read(19)

Japanese
本

Portuguese
livro

Input Comment

Post

Comment

The kanji "book" is a character that has a short horizontal line slightly below the center line of the kanji " tree ", meaning the 06/25/19 root of the tree.

Fig. 1 Screenshot from the SCROLL system showing a log inserted by a user

– Users that have English as a first Language.
– Users that have used SCROLL for more than three months.
– Users that have input more than 50 logs.

The chosen criteria allows for the analysis of the activity of users who used SCROLL regularly for an extended period of time.

Through an initial examination of the logs of the selected users, we found that the users' logs contain words that are semantically similar. Certain words belonged to a general vocabulary present throughout all the users logs, while other words belonged to a situated vocabulary present in logs created within a short and limited period of time.

– General vocabulary: learners tend to have several words throughout their logs that belong to the same semantic fields. Those words constitute general vocabulary in the users' digital trace. The general vocabulary may be related to the users' interests, fields of study or work. E.g.: botanic, animals, computer-science, etc.
– Situated vocabulary: Learners tend to save words belonging to the same semantic field within a short and limited period of time. E.g.: 12:55 pm, cat; 12:56 pm: dog; 12:56 pm: hamster; 12:57 pm: bird.

Figure 3 is an example of the presence of the situated and general vocabulary within a section of the vocabulary of one of the twenty selected users. The semantic fields associated to the words are based on a thematic analysis of the vocabulary of the learner.

Supposing that we aim to identify the intended meaning of a learner's mth log. The past vocabulary of learner is composed of $m - 1$ words.

$$vocab = word_1, word_2, ..., word_{m-1}$$

Fig. 2 Log creation and past logs on SCROLL

Based on the previous observations, we divide the vocabulary of a learner (*vocab*) into two different sets: $vocab_{general}$ that includes all the words that were inserted in the system more than five minutes before the target word, and $vocab_{situated}$ that includes the words that were inserted in the system within five minutes of the insertion of the target word.

Fig. 3 General and situated vocabulary in the vocabulary of a SCROLL user

$$vocab = vocab_{general} + vocab_{situated}$$

Identification of the user's intended meaning

In this section, we present three different methods to identify a learner's intended meaning when she/he looks up a word that has multiple meanings. The methods use the learner's past logs to identify their intended meaning. For the three different methods, we suppose that a word has n meanings:

$$word = meaning_1, meaning_2, ..., meaning_n$$

*Method*1*: identification of the intended meaning of a learner based on the general vocabulary*

In this method we use the *vocab* of a learner to identify their intended meaning. We consider that the meaning that has the highest semantic similarity with the past vocabulary is the intended meaning of the learner, as follows:

$$identifiedMeaning = \max\{(semanticSimilarity(meaning_1, vocab),$$
$$semanticSimilarity(meaning_2, vocab), ...,$$
$$, semanticSimilarity(meaning_n, vocab))\}$$

Example To illustrate how *Method*1 works, let's suppose that the language learner is translating and logging the word *calf* at 15:26. In this case, the word *calf* has two different meanings:

$meaning_1$: veal
$meaning_2$: soleus

Table 1 shows the learner's vocabulary with the dates and time of input of each word:

To identify the intended meaning of the learner we compute the following semantic similarities:

$$semanticSimilarity(veal, (thigh\ knee\ elbow\ lantern$$
$$finger\ thumb\ little\ finger\ nut\ rice\ paddy\ nail\ scissors$$
$$thermometer\ toothpaste\ shaver))$$

Table 1 Learner's vocabulary and time of input

Word	Time
Thigh	2018/11/08 15:25
Knee	2018/11/08 15:24
Elbow	2018/11/08 15:18
Lantern	2018/11/08 15:08
Finger	2018/11/04 21:07
Thumb	2018/11/04 21:06
Little	2018/11/04 21:05
Finger	2018/11/04 21:05
Nut	2018/11/03 14:15
Rice	2018/11/03 14:14
Paddy	2018/11/03 14:14
Nail	2018/11/03 14:08
Scissors	2018/11/03 14:08

$$semanticSimilarity(soleus, (\textit{thigh knee elbow lantern}$$
$$\textit{finger thumb little finger nut rice paddy nail scissors}$$
$$\textit{thermometer toothpaste shaver}))$$

If the semantic similarity between the word *veal* and the past vocabulary is the highest, we consider that the learner should get a translation for *veal*, as it is his/her intended meaning. Otherwise, we consider that *soleus* is the intended meaning of the learner.

*Method*2 : *word meaning identification based on general and situated vocabulary*

Similarly to *Method*1, the intended meaning of the learner is the meaning that has the highest semantic similarity with the past vocabulary. However, $vocab_{situated}$ can provide a more precise context regarding the intended meaning of the learner. Thus, ten times more weight is given to the semantic similarity between $vocab_{situated}$ and the different meanings of the target word, compared to $vocab_{general}$ and the different meanings of the target word.

$$identifiedMeaning = \max\{(10 * semanticSimilarity(meaning_1, vocab_{situated}$$
$$+ semanticSimilarity(meaning_1, vocab_{general}),$$
$$(10 * semanticSimilarity(meaning_2, vocab_{situated}$$
$$+ semanticSimilarity(meaning_2, vocab_{general}),$$
$$(10 * semanticSimilarity(meaning_n, vocab_{situated}$$
$$+ semanticSimilarity(meaning_n, vocab_{general})))\}$$

Example To illustrate how *Method*2 works, let's suppose again that the language learner is translating and logging the word *calf* at 15:26. As stated previously, the word *calf* has two different meanings:

$meaning_1$: veal

$meaning_2$: soleus

Table 2 shows the learner's vocabulary with the dates and time of input of each word:

The words *thigh* and *knee* constitute $vocab_{situated}$ as they were inputted within five minutes of the target word *calf*. The rest of the vocabulary constitute $vocab_{general}$. To identify the intended meaning of the learner we compute the following semantic similarities:

Table 2 Learner's vocabulary and dates of input

Word	Time
Thigh	2018/11/08 15:25
Knee	2018/11/08 15:24
Elbow	2018/11/08 15:18
Lantern	2018/11/08 15:08
Finger	2018/11/04 21:07
Thumb	2018/11/04 21:06
Little	2018/11/04 21:05
Finger	2018/11/04 21:05
Nut	2018/11/03 14:15
Rice	2018/11/03 14:14
Paddy	2018/11/03 14:14
Nail	2018/11/03 14:08
Scissors	2018/11/03 14:08

$$10 * semanticSimilarity(veal, (thigh\ knee))+$$
$$semanticSimilarity(veal, (elbow\ lantern\ finger$$
$$thigh\ little\ finger\ nut\ rice\ paddy\ nail\ scissors$$
$$thermometer\ toothpaste\ shaver))$$

$$10 * semanticSimilarity(soleus, (thigh\ knee))+$$
$$semanticSimilarity(veal, (elbow\ lantern\ finger$$
$$thumb\ little\ finger\ nut\ rice\ paddy\ nail\ scissors$$
$$thermometer\ toothpaste\ shaver))$$

If the semantic similarity of the word *veal* with the past vocabulary is the highest, we consider that *veal* is the meaning the learner should get a translation for, as it is her/his intended meaning. Otherwise, we consider that *soleus* is the intended meaning of the learner.

*Method*3: word meaning identification based on previous logs and Most common definition

The semantic similarity between different meanings of the target word and the vocabulary sometimes have very similar values. In those cases, the results of the previous methods are not a strong indicator of the intended meaning of the learner. In such situations, we propose considering that the most common meaning of the target word is the intended meaning of the learner. The semantic similarities are considered similar if their ratio varies between 0.8 and 1. In the cases where the ratio of the semantic similarities is lower than 0.8, we use the general and situated vocabulary to identify the intended meaning of the learner as shown in *Method*2. To identify the most common meaning, we use the first entry of the New Oxford American Dictionary. The entries of each word in the New Oxford American Dictionary are ordered by meaning frequency, i.e., how common the meaning is in today's English.

`if` $semanticSimilarity(meaning_i, vocab)/semanticSimilarity(meaning_j, vocab)$
$<= 0.8$
`then` $identifiedMeaning = MostcommonMeaning$
`else`

$$identifiedMeaning = \max\{(10 * semanticSimilarity(meaning_1, vocab_{situated}$$
$$+semanticSimilarity(meaning_1, vocab_{general}),$$
$$(10 * semanticSimilarity(meaning_2, vocab_{situated}$$
$$+semanticSimilarity(meaning_2, vocab_{general}),$$
$$(10 * semanticSimilarity(meaning_n, vocab_{situated}$$
$$+semanticSimilarity(meaning_n, vocab_{general})))\}$$

Results

Setup of the evaluation

More than 40% of English words are polysemous words. In order to restrict the number of words, we limit our evaluation to homographs. Homographs are words that look similar but have different origins and different meanings. The difference between polysemous words and homographs is subtle. Lexicographers define polysemous words within

a single dictionary entry, numbering different meanings, while homographs are treated in separate dictionary entries.

Our list of homographs is based on Stork's list of homographs from the New Oxford English Dictionary (Stork 1993). We confirm the homographs on the list by checking if each of the words has at least two different entries in the Oxford American Writer's Thesaurus. The words that have two or more entries constitute our final list of homographs.

In order to identify the intended meaning of the learner, we first need to compile a list of possible meanings. For each homograph, the list of possible meanings was compiled as a list of synonyms representing its different possible meanings, e.g., fan: meaning 1: ventilator; meaning 2: admirer.

We select from SCROLL the logs that contain an English homograph, as well as an image associated to them. This selection results in 148 logs. The 148 logs containing homographs belong to 78 different users. The semantic similarity is computed between each meaning of the homograph and the past vocabulary of the user that input the homograph. The past vocabulary of each author of the homographs contains an average of 127 logs. The total number of logs used to disambiguate the homographs is 9906. The semantic similarity is calculated using the cosine similarity of the open-source toolkit Gensim (Řehůřek & Sojka, 2010). We train a Wikipedia-based word embedding using English Wikipedia with Gensim Word2Vec tool. We run the algorithm for the three different methods and obtain the identified meaning for each homograph. To evaluate if our identified meaning is in fact the intended meaning of the learner, we compare it to the image uploaded by the learner. The comparison is manual, and if the identified meaning corresponds to the image uploaded by the learner, the meaning identification is considered successful. In the cases where the image does not match the identified meaning, we consider that the meaning disambiguation is a failure.

For further evaluation of our results, we compare them to Google Translate as it is the most used language learning tool. However, Google Translate results differ depending on the chosen pair of languages and the proposed methods are independent of the target language. To tackle the issue, we chose to compare our results to those of Google Translate for the English-Japanese, English-Arabic, and English-French language pairs. Arabic, French and Japanese belong to different language families. Moreover, French and English belong to the same family of Indo-European languages. This diversity in the target languages could provide a better understanding of the success rates of the proposed methods compared to Google Translate. We translate the same set of logs to Japanese, Arabic and French. If the translation provided by Google Translate corresponds to the intended meaning of the learner, we consider the translation successful. If not, the translation is considered a failure. Speakers of Japanese, Arabic, and French identified the meaning provided by Google Translate and reported whether it corresponds to the image uploaded by the learner.

Results of the evaluation

$Method1, Method2, Method3$ result in 72.180%, 75.630%, and 83.050% of correctly identified meanings respectively. We compare the results of the proposed methods to Google Translate results on the same set of logs. As Google Translate results differ depending on the chosen pair of languages, we compare our results to those of Google Translate for the

English-Japanese, English-Arabic, and English-French language pairs. Google Translate provides an identification success rate of 75.626% from English to Japanese, 71.428% from English to Arabic, and 88.721% from English to French. Table 3 provides a summary of the results. Moreover, a chi-square test was conducted to examine whether the improvement between Method 1 and Method 3 is significant. The chi-square test showed that the difference between Method 1 and Method 3 is significant $X^2(1, N=148)=4.068, p=.043702$.

Analysis of the results

Methods 1, 2 and 3 provided a success rate that is superior to the results of Google Translate with the same set of logs from English to Arabic. Methods 2 and 3 surpassed the results of Google Translate from English to Japanese. However, Google Translate surpassed the three methods when translating the provided set of words from English to French. This could be due to the fact that both English and French are Indo-European languages and contain numerous cognates (Taylor, 1976), whereas Arabic belongs to the Afro-Asiatic language family and Japanese to the Japonic language family (Ethnologue.com). This shows that providing translation based on a language learner's past vocabulary can potentially be language independent, i.e., it is not influenced by how two different languages are related.

To have a deeper understanding of the situations in which the three proposed methods failed to identify the intended meaning of the learner, we collected the list of unsuccessful identifications, and grouped them into the following categories:

– Past vocabulary semantically closer to unintended meaning: Methods 1, 2, and 3 failed to identify the intended meaning of the learner when the past vocabulary was semantically closer to an unintended meaning. This category is the most general one as well as the most populated one.

  Example

  Target word: *fan*

  Meanings of the word: *admirer, ventilator*

  Past vocabulary of the learner: *concert, singer, song, poster*

  Intended meaning identification with proposed methods: *admirer*

  Picture uploaded by the learner representing their intended meaning: *ventilator*

  Reason of failure: The past vocabulary of the learner indicates that the intended meaning of the learner is the meaning of *admirer* whereas it is the meaning of *ventilator*.

– Different definitions of a word are semantically close: In a small number of cases, the different meanings of the target word were semantically close to each other, resulting in a misidentification of the intended meaning of the learner.

  Example

Table 3 Evaluation of the methods and comparison with Google Translate on the same set of logs

	*Method*1	*Method*2	*Method*3	Google Translate En-Jp	Google Translate En-Ar	Google Translate En-Fr
Identification rate	72.180%	75.630%	83.050%	75.626%	71.428%	88.721%

Target word: *book*

Meanings of the word: *notebook, textbook, reserve*

Past vocabulary of the learner: *write, study, read, school.*

Intended meaning identification with proposed methods: *notebook*

Picture uploaded by the learner representing their intended meaning: *textbook*

Reason of failure: The past vocabulary of the learner is semantically close to two of the meanings of the word *book*: *notebook* and *textbook*. The identified meaning was *notebook* as it had a slightly higher score with the proposed methods. However, the actual intended meaning of the learner was *textbook.*

– Only the situated vocabulary is semantically close to the intended meaning: In some cases, the overall past vocabulary of the learner was semantically closer to the unintended meaning, whereas the situated vocabulary was semantically closer to the intended meaning of the learner. This case led to a failure of identification in *Method*1 only, as the issue was solved in Methods 2 and 3.

    Example

    Target word: *fan*

    Meanings of the word: *admirer, ventilator*

    Past vocabulary of the learner: 12:01 *concert*, 12:03 *singer*, 15:09 *song*, 15:09 *actor*, 15:10 *movie*, 15:54 *poster*, 19:32 *hot*, 19:33 *summer*, 19: 34 *wind*, 19:35 *blow*

    Intended meaning identification with proposed methods: *admirer*

    Picture uploaded by the learner representing their intended meaning: *ventilator*

    Reason of failure: The general vocabulary of the learner (until 15:54) is semantically closer to the meaning of *admirer* whereas the situated vocabulary is closer to the meaning of *ventilator*, which is the intended meaning of the learner. As the general vocabulary is bigger than the situated vocabulary, *Method*1 misidentified the intended meaning. This kind of misidentification was eliminated from *Method*2 and *Method*3 by giving more weight to the situated vocabulary.

– Few past vocabulary: In some cases, the past vocabulary of the learner was formed by very few words (one to five), and did not contain any particular pattern.

    Example

    Target word: *fan*

    Meanings of the word: *admirer, ventilator*

    Past vocabulary of the learner: *water*

    Intended meaning identification with proposed methods: *admirer*

    Picture uploaded by the learner representing their intended meaning: *ventilator*

    Reason of failure: The past vocabulary is too small and contains words unrelated to both meanings of the target word

– Situated vocabulary semantically closer to unintended meaning: In some cases, the general past vocabulary of the learner was semantically closer to the intended meaning, whereas the situated vocabulary was strongly closer semantically to the unintended meaning. This issue led to identification failures in the cases of *Method*2 and *Method*3.

    Example

    Target word: *fan*

    Meanings of the word: *admirer, ventilator*

    Past vocabulary of the learner: 19:32 *hot*, 19:33 *summer*, 19: 34 *wind*, 19:35 *blow*, 19:55

swing, 20:21 *blades*, 20:30 *folded*, 20:40 *concert*, 21:41 *singer*, 21:41 *song*, 21:42 *actor*
Intended meaning identification with proposed methods: *admirer*
Picture uploaded by the learner representing their intended meaning: *ventilator*
Reason of failure: The situated vocabulary of the learner (from 21:41 to 12:42) is semantically closer to the meaning of *admirer* whereas the general vocabulary is closer to the meaning of *ventilator*, which is the learner's intended meaning. As the situated vocabulary has a bigger weight than the general vocabulary in *Method*2 and *Method*3, the intended meaning was misidentified. This kind of misidentification only occurred in *Method*2 and *Method*3.

– The most common meaning is not the intended meaning: *Method*3 failed to identify the intended meaning in some cases when the different meanings of the target word had similar semantic similarity with the past vocabulary, and the most common meaning of the word happened to be the unintended meaning of the learner.

Example
Target word: *calf*
Meanings of the word: *veal, soleus*
Past vocabulary of the learner: *animal, knee, leg, cow*
Intended meaning identification with proposed methods: *veal*
Picture uploaded by the learner representing their intended meaning: *soleus*
Reason of failure: The past vocabulary is as semantically close to both meanings of the word *veal*. In those cases, *Method*3 selects the most common meaning. In this case, the most common meaning of the word *calf* is the meaning of *veal*. However, the learner uploaded a picture that represents the *soleus*.

Figure 4 shows the reasons of failure of the different methods as well as their distribution.

Fig. 4 Reasons of failure of Methods 1, 2 and 3
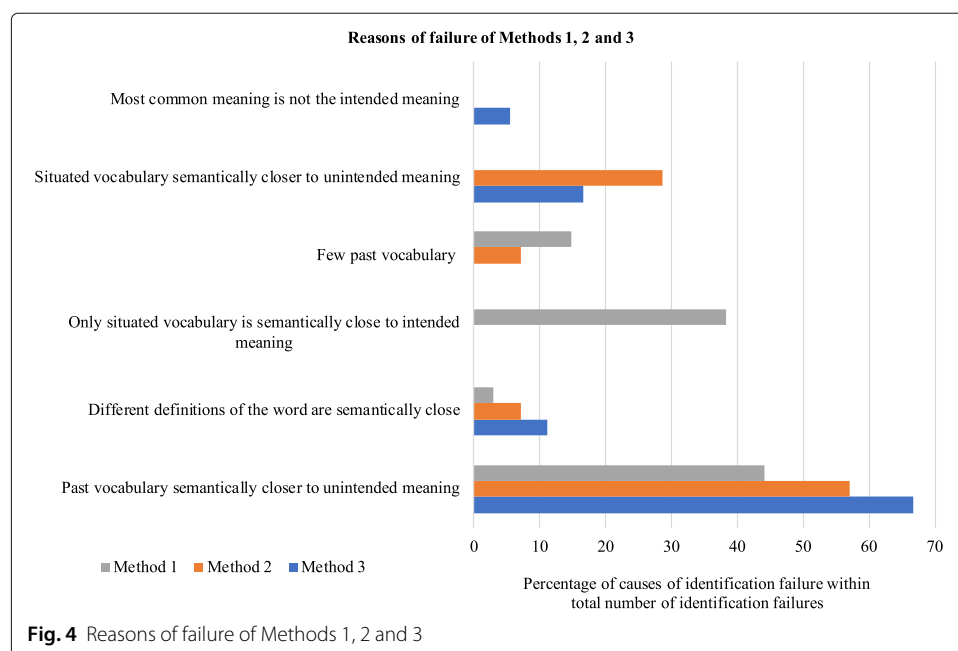
Discussion

This paper shows the potential that informal ubiquitous vocabulary learning environments have to provide language learners with personalized translations based on their learning activity. Learners choose the words they want to learn. Their choice of words does not come from a list in a textbook, but from the learner's interests, activities and surroundings. We proposed methods to identify the learners' intended meaning when they look up a polysemous word or homograph. If applied, those methods can lead to a personalized and contextualized translation that is beneficial to the learning process.

Considerations and opportunities Learning the correct translation in the intended context of use could improve the learner's communication skills or reading comprehension skills. However, even though the intended meaning of the learner is identified and shown to them, the learner should still be exposed to all the different meanings of a word, and be aware of the existence of different meanings in different contexts of use. Highlighting the intended meaning would help the learner memorize and connect the translation to their current context.

If the language learner is following language classes with the help of an instructor, word sense disambiguation could help the language teacher track writing mistakes based on the intended meaning of the learner. For example, in situations where students submit written compositions through a Learning Management System (LMS), the LMS could be linked to their learning logs. The teacher could then view not only the written composition but also the words that the learner looked up and used in their writing, as well as the intended meaning of the student in the cases of polysemous words. Understanding the learner's intended meaning would give the teacher the opportunity to check if the usage of the word was correct in the sentence and give feedback on a better formulation based on the intended meaning. Moreover, the teacher could then give detailed explanations on the most appropriate form of usage in the particular context intended by the student.

Moreover, numerous studies have shown a correlation between trust in and usage of a system (Lee & Moray, 1994; Muir & Moray, 1996). Privacy concerns and uncertainties negatively affect the usage of a system (Zhou, 2012). SCROLL's users did not enable the location on 65.7% of SCROLL's logs. Enabling the location on SCROLL does not require much effort as SCROLL displays a notification upon login asking the user for permission to enable the location. Moreover, enabling the location in SCROLL would offer the users benefits like location-based word recommendations. Even though enabling the location services is easy and useful, many users choose not to do it, implying they may lack trust in the system. The methods proposed identify the intended meaning of the learner using their past activity and might create privacy concerns. To maintain the trust of the user, the system has to maintain transparency and the user has to be aware of the method behind the word meaning disambiguation and give consent to use their past activity for a personalized translation.

Limitations The number of logs analyzed is quite small as we had three main restrictions: the logs had to contain an English word, the word had to be a homograph, and the learner had to have an image uploaded in order to confirm our meaning identification. However, the results of the methods clearly show that the past activity of a learner in an informal ubiquitous learning environment is a good indicator of their intended meaning.

In the proposed methods, we considered that the situated vocabulary is the vocabulary saved within a short period of time of the target word, giving it a temporal definition. SCROLL in particular, and informal ubiquitous language learning environments in general, allow the capture of the user's location while logging a new word. Another option would be to use the location to define what constitutes the situated vocabulary. However, in our available data, many users did not allow the system to access their location, others saved improbable locations (e.g., "the sea"), and some users were continuously using the system from only one location. As the available data represents real life situations and challenges, we considered that a temporal division would be more reliable in identifying the intended meaning of the learner. A temporal definition of the situated vocabulary can represent the spatial aspect when the user does not permit access to their location. Moreover, the temporal definition includes situations in which the user changes activities or studies different subjects without moving locations.

An important limitation of this paper lies in the evaluation itself. To evaluate the success or failure of the intended meaning identification, we rely on the images uploaded by the learner. However, the images uploaded by the users represent objects or places, as it is difficult to find or take pictures representing abstract concepts. This factor resulted in a selection of logs, where the intended meaning of the learner relates mostly to objects. This limitation does not deny the potential shown by the results in identifying the intended meaning based on past vocabulary. However, a different evaluation should be designed to include non-material concepts in the set of logs studied.

A failure to identify a leaner's intended meaning occurred in cases in which the users of SCROLL had no or very few past vocabulary words saved on the system. *Method*3 tackles part of the problem by selecting the most common definition of a word when different definitions have similar semantic similarities with the past vocabulary. Thus, *Method*3 allows the selection of the most common definition in cases in which there is no previous vocabulary uploaded by the learner. However, *Method*3 failed to predict the intended meaning when the learner had saved only a few words previously, and those words happened to be significantly more semantically similar to the unintended definition of the target word. It would be important to determine the size of the vocabulary needed in order to have a reliable outcome when applying the method.

Another limitation is imposed by the different meanings we chose for the isolated words. In this paper, we selected the first synonym for each meaning in the Oxford American Writer's Thesaurus. We then computed the semantic similarity between the synonym and the past vocabulary of the learner. Choosing different synonyms might have led to different semantic similarities. Another alternative would be to consider the whole definition of the word as a meaning, and compute the semantic similarity between the definition and the past vocabulary of the learner.

Homographs have different meanings that are clearly not related to each other. A typical example is the word *bank*, with one meaning being *financial institution* and the other referring to the *side of a river*. However, the meanings of a polysemous word have a common origin and some of them are highly semantically related (Lehrer, 1990). For example, the word *book* could mean a *printedwork* or a *bound set of blank sheets for writing in*. In the work, we considered homographs to test the method and this allowed us to avoid situations where the meanings were very similar to each other and belonged to the same semantic fields. It would be more difficult to pinpoint the exact meaning the learner is

looking for if we had also considered polysemous words. However, this could be solved by providing the learner all of the definitions that could correspond to their intended meaning, eliminating meanings that are very semantically far from the intended meaning. Providing different definitions would still teach the learner the meaning(s) of a word in its intended context of use.

Conclusion

In this work, we aimed to identify the intended meaning of language learners when they look up a polysemous word or a homograph in a distance ubiquitous language learning environment. Following observations on learner's logs in a ubiquitous language learning environment, we conclude that learners tend to have a general vocabulary, i.e., words that are semantically related throughout their vocabulary, as well as a situated vocabulary, i.e., words that are semantically related within a short period of time. Based on those observations, we proposed three methods, that use their past vocabulary to identify their intended meaning. The first method considers that the intended meaning of the learner is the one that is the most semantically similar to the learner's past vocabulary. The second method builds on the first method but gives more weight to the vocabulary that the learner logged shortly before the target word. The third method addresses situations where the semantic similarities between the different meanings of the word and the past vocabulary have similar values. In those cases, the method considers that the intended meaning of the learner is the most common meaning in the target language. The three methods were evaluated using 148 logs of SCROLL, a ubiquitous informal language learning environment. The success rates of the three methods were respectively 72.180%, 75.630%, and 83.050%. This work shows that the past activity of language learners in informal ubiquitous language learning environments could be used to identify their intended meaning when learning a new word.

Author details
[1] Kyoto University, Department of Social Informatics, Graduate School of Informatics, Kyoto, Japan. [2] Academic Center for Computing and Media Studies, Kyoto University, Kyoto, Japan.

References
B. T. Atkins, K. Varantola, Monitoring dictionary use. International Journal of Lexicography. **10**(1), 1–45 (1997)
M. Azab, A. Salama, K. Oflazer, H. Shima, J. Araki, T. Mitamura, An nlp-based reading tool for aiding non-native english readers (2013)
A. Boulton, S. De Cock, Dictionaries as aids for language learning. International Handbook of Modern Lexis and Lexicography, 1–17 (2017). https://doi.org/10.1007/978-3-642-45369-4_25-1
P. De Brabanter, M. Kissine, *Utterance Interpretation and Cognitive Models. vol. 20.* (Brill, 2009). https://doi.org/10.1163/9789004253148

V. Demouy, A. Jones, Q. Kan, A. Kukulska-Hulme, A. Eardley, Why and how do distance learners use mobile devices for language learning?. The EuroCALL Review. **24**(1), 10–24 (2016)

D. N. Dimakopoulos, G. D. Magoulas, in *Innovative Mobile Learning: Techniques and Technologies*, An architecture for a personalized mobile environment to facilitate contextual lifelong learning (IGI Global, 2009), pp. 232–254. https://doi.org/10.4018/978-1-60566-062-2.ch012

C. Ducar, D. H. Schocket, Machine translation and the l2 classroom: Pedagogical solutions for making peace with google translate. Foreign Language Annals. **51**(4), 779–795 (2018)

S. Eom, M. Dickinson, R. Sachs, in *Proceedings of the Seventh Workshop on Building Educational Applications Using NLP*, Sense-specific lexical information for reading assistance (Association for Computational Linguistics, 2012), pp. 316–325

Ethnologue.com, Language Families. https://www.ethnologue.com/browse/families. Accessed 24 June 2019

C. Gurrin, A. F. Smeaton, A. R. Doherty, Lifelogging: Personal big data. Foundations and Trends® in information retrieval. **8**(1), 1–125 (2014)

M. N. Hasnine, K. Mouri, B. Flanagan, G. Akcapinar, N. Uosaki, H. Ogata, in *Proceedings of the 26th International Conference on Computer in Education*, Image recommendation for informal vocabulary learning in a context-aware learning environment, (2018), pp. 669–674

C. R. Heil, J. S. Wu, J. J. Lee, T. Schmidt, A review of mobile language learning applications: Trends, challenges, and opportunities. The EuroCALL Review. **24**(2), 32–50 (2016)

H. Hilton, The link between vocabulary knowledge and spoken l2 fluency. Language Learning Journal. **36**(2), 153–166 (2008)

L. Jin, E. Deifell, Foreign language learners' use and perception of online dictionaries: A survey study. Journal of Online Learning and Teaching. **9**(4), 515 (2013)

P. Joyce, L2 vocabulary learning and testing: The use of l1 translation versus l2 definition. The Language Learning Journal. **46**(3), 217–227 (2018)

B. Laufer, P. Nation, Vocabulary size and use: Lexical richness in l2 written production. Applied linguistics. **16**(3), 307–322 (1995)

J. D. Lee, N. Moray, Trust, self-confidence, and operators' adaptation to automation. International journal of human-computer studies. **40**(1), 153–184 (1994)

A. Lehrer, Polysemy, conventionality, and the structure of the lexicon. Cognitive Linguistics (includes Cognitive Linguistic Bibliography). **1**(2), 207–246 (1990)

G. A. Miller, P. M. Gildea, How children learn words. Sci Am. **257**(3), 94–99 (1987)

E. E. Mueller, The Elephant in the Room: Machine Translation in the Language Classroom (2013). https://urlzs.com/mxtPm. Accessed 24 Apr 2018

B. M. Muir, N. Moray, Trust in automation. part ii. experimental studies of trust and human intervention in a process control simulation. Ergonomics. **39**(3), 429–460 (1996)

W. E. Nagy, On the role of context in first-and second-language vocabulary learning (1995). Technical report, Champaign, Ill.: University of Illinois at Urbana-Champaign, Center for the . . .

H. Ogata, M. Li, B. Hou, N. Uosaki, M. M. EL-Bishouty, Y. Yano, Scroll: Supporting to share and reuse ubiquitous learning log in the context of language learning. Research & Practice in Technology Enhanced Learning. **6**(2) (2011)

H. Ogata, N. Uosaki, K. Mouri, M. Hasnine, V. Abou-Khalil, B. Flanagan, in *Workshop Proceedings of the 26th International Conference on Computer in Education*, Scroll dataset in the context of ubiquitous language learning, (2018), pp. 418–423

R. Řehůřek, P. Sojka, in *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, Software Framework for Topic Modelling with Large Corpora (ELRA, Valletta, 2010), pp. 45–50. http://is.muni.cz/publication/884893/en

M. Sharples, J. Taylor, G. Vavoula, in *Proceedings of mLearn, vol. 1*, Towards a theory of mobile learning, (2005), pp. 1–9

L. S. Stæhr, Vocabulary size and the skills of listening, reading and writing. Language Learning Journal. **36**(2), 139–152 (2008)

D. G. Stork, A complete list of homographs from the oxford english dictionary. American Speech. **68**(3), 330–334 (1993)

I. Taylor, Similarity between french and english words—a factor to be considered in bilingual language behavior?. Journal of Psycholinguistic Research. **5**(1), 85–94 (1976)

T. Zhou, Examining location-based services usage from the perspectives of unified theory of acceptance and use of technology and privacy risk. Journal of Electronic Commerce Research. **13**(2), 135 (2012)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.