

**Dual Use of Linguistic Resources:
Evaluation of MT Systems and Language Learners**

Lisa Decrozant
University of Maryland, Dept. of French and Italian
Army Research Laboratory
decrozan@arl.mil

Clare R. Voss
Army Research Laboratory
voss@arl.mil

Please send correspondence to:

Lisa Decrozant
Attn. AMSRL-IS-CI
2800 Powder Mill Rd.
Adelphi, MD 20783-1145

Dual Use of Linguistic Resources: Evaluation of MT Systems and Language Learners

1. Introduction

Human translators working with “embedded machine translation (MT) systems”¹ on the task of filtering text documents in a foreign language often have limited training in the foreign language they encounter. For our MT system users who are also language learners, we are developing a suite of linguistic tools that enable them, on the same laptop platform, to perform their foreign language filtering tasks using a combination of OCR, MT, IR and language sustainment tools.² Thus we have begun constructing linguistic test suites that can serve the dual needs we have for the evaluation of MT systems and language learning.³

In this paper, we present our pilot work (i) defining and constructing a semantic domain of spatial expressions as a test suite, (ii) testing our MT system on the translations of these expressions, and (iii) testing language learners’ ability to translate these expressions. Our results show that, for English-to-French translation of a small set of spatial expressions, neither a commercially viable MT system nor intermediate-level students are adequately trained to identify explicit and implicit (ambiguous) paths of motion.

¹ The term *embedded MT system* we adopt from Voss and Reeder, 1998.

² We are creating a single interface for the MT system and the language sustainment tools that enables users to guide their own learning during MT-aided tasks, such as filtering, in contrast to single-purpose tutoring systems (e.g., Holland et al., 1995)

³ For others addressing multiple uses of linguistic resources, see NLP-IA (1998).

2. Identifying Linguistic Issues for Evaluation

English and French are known to “diverge” in their expression of spatial relations⁴: given a spatial expression in one of these languages, the process of translating it will fail if a simple word-for-word replacement strategy is used, whether translated by an MT system or by language learners.

2.1 Directional Particles & Prepositions

Klippel (1992) documents a divergence that exists between English and French in the semantics of direction. In English, directional particles, such as ‘*up*’ and ‘*down*’, may appear following a verb of motion, giving the verb’s event a directed motion reading. In French, however, there are no equivalent lexical items corresponding to these English directional particles. Instead, the semantics of direction is expressed elsewhere. Klippel also observes more generally, following Talmy (1983), that directions are typically incorporated within the French motion verb. Example 1 shows one such case where the English verb-plus-preposition ‘*went up*’ translates into the French verb ‘*est monté*’ without a preposition.

1. E⁵ The child **went up** the stairs.
F: L’enfant est monté l’escalier.
g: the child ascended the stairs

2.2 Locational Prepositions

English and French also diverge in their expression of spatial relations with respect to a second group of prepositions. As noted by Jackendoff (1983), English

⁴ We use the term *divergence* as in Dorr, 1994.

⁵ In the examples, ‘E’=English, ‘F’=French, ‘g’=Gloss, ‘p’=paraphrase

locational, or Place-type prepositions may be ambiguous because they may also have a directional, Path-type reading.⁶ For example, sentence 2. is ambiguous in English. In the 2(a) reading, the bottle moves along a path as it floats, starting away from the bridge and ending up under the bridge. In the 2(b) reading, the bottle remains under the bridge, as it floats: no path is specified, only the general location of where the floating took place. In French, however, the word-equivalent of a locational preposition typically has only the 2(b) locational reading, not the (2a) directional reading.

-
2. E: The bottle floated **under** the bridge.
2a. p: the bottle floated to a place under the bridge
2b. p: the bottle floated while under the bridge
F: La bouteille a flotté sous le pont. (sense 2b)
g: the bottle floated under the bridge
-

We selected the domain spatial expressions for evaluation in part because, as example 3 shows, the ambiguity of English spatial prepositions may significantly interfere with the task of accurate message understanding—whether by MT systems or second language learners. As Taylor and White (1998) point out, in a real-world, task-based evaluation of MT systems or language learners, it is the correct and incorrect consequences of our users’ actions based on their understanding of a foreign language text document that is the measure of interest. Such measures of effectiveness are difficult to obtain and researchers, outside of the field, must rely instead on linguistically based measures of performance. Thus, our approach has been to build our test suite relying on extensive pre-existing, linguistically motivated spatial language research (e.g.,

⁶ In a pilot study, three native English speakers we tested consistently identified 35 locational prepositions in English with this form of ambiguity.

Herskovits, 1986; Jackendoff, 1983; Lindstromberg, 1998; Olivier and Gapp, 1998; Vandeloise, 1992;).

Example 3 is a linguistically simple variation on example 2: both have manner of motion verbs (*float*, *march*) and locational prepositions (*under*, *in*). In each case the prepositional phrase (PP) may be an argument to the verb (the 2a. and 3a. paraphrases) or an adjunct (the 2b. and 3b. paraphrases). Notice that, if the exact location of the troops were mission-critical information, then this ambiguity should not go undetected. In one case, the troops have changed locations by moving into the canyon, while in the other, the troops are remaining in the canyon.

-
3. E: The troops marched **in** the canyon.
3a. p: The troops entered the canyon marching
3b. p: The troops were marching about in the canyon
-

3. Developing Linguistic Resources for Evaluation

In order to assess how accurately and consistently MT systems handle spatial language and how effectively second language learners are being taught about spatial language, we followed these steps in constructing a spatial expressions evaluation dataset. First we built a masterlist of English prepositions from several sources (Lindstromberg, 1998; Websters, WWWsite) and then created a sublist of only spatial prepositions, based on the judgments of three native English speakers, two of whom were linguistically trained and one who was not.

Second, we constructed English sentences where the spatial PP was systematically composed into different syntactic positions, as shown in Figure 1. This enabled us to examine the range of Path/Place-type ambiguity in the resulting spatial expressions. The spatial

prepositions were placed in contexts where only a place or a path interpretation was feasible, as well as in contexts where the reading was ambiguous. For example, PP's with the preposition 'across' were composed in (i) a verb's subcategorized argument for the verbs 'live' and 'dance', as in 'he lived/danced [PP across the street]', (ii) as a verb's non-subcategorized argument for the verbs 'scare' and 'sneeze', as in 'he scared the child [PP across the street]' and 'he sneezed the cards [PP across the table]', and (iii) as an adjunct outside the VP for the verb 'eat', as in 'he ate dinner [PP across the street]'.

subcat Arg	nonsubcat Arg	adjunct
"He lives across the street."	"He sneezed the cards across the table."	"He ate dinner across the street."
"He danced across the street."	"He scared the child across the street."	

Figure 1, a row of ESE data set, with preposition "across"

Once this English spatial expression (ESE) dataset was constructed, our third step was to elicit translations of a subset of these sentences into French. Our translator was a native French speaker, fluent in English, with a PhD. in linguistics from a university in the U.S. Our translator's extensive training in linguistics made it possible for us to be quite specific about the English ambiguities present in the sentences that we needed him to translate.

Our efforts have yielded the following datasets: an English preposition list, an English spatial preposition list, a dataset of English spatial expressions sorted by their spatial preposition and syntactic structure, and a dataset of high-quality French translations of a proper subset of the ESE dataset.

4. Using Linguistic Resources to Evaluate an MT System

One of the objectives of our work is to support users of the embedded MT systems that our laboratory has been involved in developing. These systems were designed to be 'good enough' for filtering or relevance analysis of hard-copy, open source text documents.⁷ The ESE dataset was developed as part of an ongoing effort to expand our evaluation test suites. Here we report on a preliminary test that explored the feasibility of using sentences from the ESE dataset with their human translation into French, to evaluate one MT engine that we know is being used in the field.

Eight sentences from the ESE dataset were selected and run through the MT engine from English to French. The results of these automatic translations were then compared to the human translator's translations. Two groups of prepositions, corresponding to the two types of divergences discussed above, were of interest to us.

4.1 Default Place Readings

First, we were curious about how ambiguous path/place readings were handled, given that the MT engine we were working with was designed to produce only one, preferred translation per input sentence, as is common for commercial MT products. We predicted that only the place reading would appear in the French MT results. We knew from discussions with MT developers that they rely heavily on hand-coded dictionaries in creating their on-line lexicons. Since English and English-French dictionaries list locational prepositions, such as those in examples 2 and 3, with only a place

⁷ Church and Hovy (1993) spelled out this notion of 'good enough' MT and Resnik (1997) has introduced a clever method to test this.

reading, not a path reading, it seemed most likely that only the place reading would appear in the French MT results.

Another reason we expected place readings for the ambiguous phrases was that they are the direct result of the shortest path through an MT system, that is, with simple word replacements. Our predictions proved correct. Five sentences were ambiguous with both place and path readings, but all received only a place reading in the MT translations:

Test Sentences	MT Output
1. He danced behind the screen.	1. Il a dansé derrière l'écran.
2. He carried his luggage in the airplane.	2. Il a porté son bagage dans l'avion.
3. He carried his luggage inside the restaurant.	3. Il a porté son bagage à l'intérieur du restaurant.
4. He jumped on the bed.	4. Il a sauté sur le lit.
5. They danced in the room.	5. Ils ont dansé dans la chambre.

These results led us to predict that in our sentence ‘*The troops marched in the canyon*’, the MT engine would produce only the translation that meant the troops were marching while remaining in the canyon. This was indeed what the engine did produce when we tested it.

4.2 True Path Readings

Second, we wanted to see what happened to the unambiguous path readings, given that the MT engine needed only a lexical pattern recognition to detect the English verb-preposition combination and then follow the well-documented conversion to French (Dorr, 1994). As shown in example 4, the English spatial semantics is redistributed: the manner of motion in the main verb is moved out to an adjunct in the French (*en marchant*), while the motion of going into the canyon is lexicalized in the main French verb and preposition (*entrer* and *dans*).

4. E: The troops marched into the canyon.
F: Les soldats sont entrés dans la gorge en marchant.
g: The troops entered in the canyon marching.

We suspected however that the unambiguous path readings might not be properly detected, given the English-French divergence with respect to directional particles and prepositions discussed above.

Test Sentences	MT Output
1. He carried his luggage across the street.	1. Il a porté son bagage à travers la rue.
2. He climbed down the mountain.	2. Il s'est élevé en bas de la montagne.
3. The woman jumped out of the cake.	3. La femme a sauté du gâteau. ⁸

Our suspicions were correct; the MT engine did not correctly translate the three unambiguous path-only readings we tested. Surprisingly, the actual MT-generated translations failed to capture *any* path interpretation at all. Example 5 shows that the MT system again produced the direct result of the shortest path through an MT system, with simple word replacements. Since the English ‘*into*’ translated to ‘*dans*’, the overall result was incorrect: the translation produced the unambiguous French place-only reading.

5 E: The troops marched into the canyon.
MT-F output:
Les soldats ont marché dans la gorge.
g: The troops marched in the canyon.

The results of the MT experiment allow us to conclude that for ‘true-path’ pattern sentences, the MT system will most likely fail to output an accurate translation.

Our predictions for the behavior of the first group of prepositions proved correct. On the second group of prepositions, we predicted accurately that the MT engine would not produce the correct translation, however, we failed to predict the specific translations that were output. The MT

⁸ Although technically correct, this translation is the result of a “simple word replacement” strategy on the part of the MT system, and not a sophisticated translation using semantic interpretation.

engine we are working with allows users to create their own lexicon entries that supercede those of the built-in general-purpose system lexicon. Our next steps will be to test other prepositions and to examine how the lexicon entries we create will alter the translations.

5. Using Linguistic Resources to Evaluate Language Learners

We are interested in the idea that learners can benefit from viewing parallel sentence-aligned text, as has been explored for cross-training of French-speakers learning Haitian Creole (Rincher, 1986). We would expect that divergences are readily understood by language learners when presented with parallel text. Our first step however, before exploring this possibility for teaching, has been to use the ESE dataset to evaluate second language learners to determine if they encounter the problems with spatial language as MT system did.

Fourteen intermediate-level, French language learners were given the same sentences from the data set used in the MT pilot experiment and were asked to translate into French. They were told explicitly that some of the sentences might be ambiguous. They were also given a spatial expression that was ambiguous as an example and the two interpretations of that expression were explained with paraphrases.

Student	True-Path	Default-Place
#1	1/3	0/5
#2	1/3	0/5
#3	1/3	0/5
#4	1/3	3/5
#5	1/3	2/5
#6	1/3	0/5
#7	0/3	4/5
#8	0/3	0/5
#9	2/3	2/5
#10	1/3	2/5
#11	1/3	2/5
#12	2/3	2/5
#13	1/3	5/5
#14	1/3	5/5

Due to their level of French, the college students were not always aware of the divergence in the expression of spatial paths. When faced with unambiguous path sentences (“true path” column in data table), the majority gave a simple word replacement translation, just as we had found in the MT system output. None of the students were able to correctly translate all three test sentences.

In contrast to this, when translating into French the English sentences with default place-type prepositions (“default place” column in data table), a few students were able to consistently incorporate the spatial meaning of the English preposition into the French verb and properly disambiguate the test sentences. Nonetheless, these students were not able to use this knowledge in their translations of the “true path” sentences.

This pilot experiment has given us a preliminary look at learners’ understanding of cross-linguistic divergences in spatial expressions. Further testing of this domain with other sentences and with more advanced students is still needed.

6. Conclusions and Future Work

We have developed a test suite of spatial expressions as part of our ongoing support work evaluating the embedded MT system prototypes and the language sustainment tools being developed in-house. The French language examples discussed above show how problematic the domain of spatial language is for both MT and for language learners.

References

K. Church and E. Hovy (1993) “Good Applications for Crummy Machine Translation.” *Machine Translation*, 8, 239-258.

B. Dorr (1994) "Machine Translation Divergences: A Formal Description and Proposed Solution", *Computational Linguistics*, 20(4), 597--633.

A. Herskovits (1986) *Language and Spatial Cognition*. Cambridge U. Press, England.

M. Holland, J. Kaplan, M. Sams eds.(1995) *Intelligent Language Tutors: Theory Shaping Technology*. Lawrence Erlbaum, Mahwah, NJ.

R. Jackendoff (1983) *Semantics and Cognition*. Cambridge, MA: MIT Press.

E. Klipple (1997) "Prepositions and Variation." In Di Sciullo (ed.) *Projections and Interface Conditions*, Oxford U. Press, NY, pp. 74-108.

S. Lindstromberg (1998) *English Prepositions Explained*. John Benjamins, Philadelphia, PA.

NLP+IA (1998) *Proceedings of the Natural Language Processing and Industrial Applications Conference*, Moncton, New Brunswick, Canada.

L. Decroquant, C.R. Voss "Cross-Linguistic Resources for MT Evaluation and Language Training"

A. Díaz de Ilarraza et al. "Integration of NLP tools in an Intelligent Computer-Assisted Language Learning Environment for Basque: IDAZKIDE"

C. Moghrabi "Preliminary Study for *Minori-Fra*, A Software for Teaching French in a Minority Setting."

P. Olivier and K-P.Gapp, eds. (1998) *Representation and Processing of Spatial Expressions*. LEA Publishers, Mahwah, NJ.

Resnik, P. (1997) "Evaluating Multilingual Gisting of Web Pages." In *Working Notes of the AAAI Spring Symposium on Natural Language Processing for the WWW*, Palo Alto, CA.

Rincher, D. (1986) *Franse ak Angle San Traka e San Dlo Nan Je: Nou Pr'al Potoprens*. Rincher and Associates, Forest Hills, NY

L. Talmy (1985) "Lexicalization Patterns: Semantic Structure in Lexical Forms." In T. Shopen (ed.) *Language Typology & Syntactic Description*. Cambridge U. Press, England, pp. 57-149.

K. Taylor and J. White (1998) "Predicting What MT is Good for: User Judgements and Task Performance" In Proceedings of the Conference of the Association for Machine Translation in the Americas (AMTA-98). Langhorne, PA.

C. Vandeloise (1991) *Spatial Prepositions*. U. of Chicago Press, Chicago, IL

Websters, <http://www.mw.com>

C. R.Voss and F. Reeder, eds. (1998) *Workshop on Embedded MT Systems: Design, Construction and Evaluation of Systems with an MT Component*. Held in conjunction with of the Conference of the Association for Machine Translation in the Americas (AMTA-98). Langhorne, PA