ReCALL **17** (1): 139–154. © 2005 Cambridge University Press
DOI: 10.1017/S0958344005001011 Printed in the United Kingdom

Content-based support for Persian learners of English

VAHID MIRZAIEAN
Iran University of Science and Technology, Tehran, Iran

ALLAN RAMSAY
Department of Computation, UMIST, PO Box 88, Manchester M60 1QD, UK
(*email: Allan.Ramsay@manchester.ac.uk*)

Abstract

People learning a new language have problems with various aspects of the language they want to learn – with its grammar, with its pronunciation, with its morphology... CALL tools have been provided for most of these linguistic levels, but comparatively little has been done about helping students with problems of meaning. We will argue that language technology is now mature enough to support tools that address at least some semantic issues, and we will describe a tool that attempts to provide appropriate support for a specific class of learners.

1 Content-based support

When someone learns a new language, they have to learn a lot of things. They have to learn what sounds to make and how to recognise sounds; they have to learn the structural (grammatical and morphological) rules of the target language; and they have to learn the relationship between ideas and their linguistic realisations in the target language – in other words, they have to learn how to express 'meanings' in the target language.

Providing tools to help students express the right meaning is, to put it mildly, a fairly tricky task. You can write computational tools to support students with structural (syntactic/morphological) aspects of language by taking programs that carry out the appropriate kind of analysis and 'relaxing' them. To help students learn about grammar, for instance, you can take a parser for the language they are trying to learn and ask it to provide a syntactic analysis of what the student has written. If what they have written does not actually conform to the grammar that your parser uses, you can relax the rules of the grammar until an analysis is found, and then report these violations of the grammar to the learner for them to reflect on.

This is a fairly straightforward idea. You need a good parser, and a good underlying

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

grammar, but these are now becoming available, at least for some languages (see Flickinger, 2000 for a recent survey of parsing algorithms). You also need to understand how to transform the rather opaque messages that come out of such systems so that the learner can understand them – lists of violated constraints using technical terminology from complex grammatical formalisms tend not to help naive users all that much. But the basic idea is fairly obvious, and there are now a number of implementations (e.g. Heift, 2003; Schulze, 1999).

Doing the same for meanings is rather more challenging. You need a set of rules for carrying out the 'appropriate kind of analysis', and you need some way of relaxing these rules. Both of these are challenging. We describe below some initial experiments in providing such a tool.

We start by considering what you could do if you did have a computer system that could construct representations of what something meant. How could you use such a tool to help someone who is learning a foreign language?

To answer this, we consider two further questions: what kinds of semantic errors do learners make, and how can you tell whether what someone says is what they meant?

1.1 What kinds of semantic errors do learners make?

Learners make a range of mistakes when trying to convey their intended messages. Some of these are straightforward lexical confusions – often 'false friends' – where the learner just uses the wrong word for the concept they want to convey. Others are more subtle. In particular, there are a number of words and constructions which have 'weak' interpretations, i.e. which do not mean all that much when seen in isolation, but which express a range of relationships between other concepts mentioned in the text. These include the following:

Tense and aspect markers: most languages provide devices for depicting the temporal structure of the events being reported. Many languages provide inflectional markers of tense (when the event happened) and aspect (whether it was completed or not, whether it happened once or several times). These markers are slippery – the same surface construction may indicate very different temporal relations when used in different contexts. The English present progressive, for instance, may denote a single incomplete event ('*I am eating it*') or an iterated instantaneous event ('*I am hitting him*') or a temporary state of affairs ('*I am living in Derbyshire*') (Moens & Steedman, 1988). Such items pick up their meanings from the surrounding context, in ways which are fairly hard to describe. They also tend to differ from language to language, and they can be hard to learn because of the difficulty that even native speakers encounter when trying to explain them.

Prepositions: like tense and aspect markers, prepositions mean very little in isolation. If you look up a word like '*in*' in a standard dictionary, you generally find something like 'preposition expressing *in*clusion or position with*in* limits of space, time, circumstance, etc.',[1] which hardly makes things clearer. Furthermore,

[1] *Pocket Oxford Dictionary of Current English*, 1970: most dictionaries, e.g. *Collins English* dictionary, 1992 , are just as circular.

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

prepositions also mirror tense/aspect markers in changing their effects depending on the surrounding linguistic context: the various meanings of '*in*' in '*I'll arrive in London in a car in a bad temper in ten minutes*' are extremely diverse, with not much more than a metaphorical connection between them.

Reference: discourse is about entities and the relationships between them. Some of the entities you need to talk about are already highly salient in the context, some are known to both parties but have not been mentioned recently, some are new to the conversation. Again the markers used in different languages for distinguishing between different ways of introducing and referring to entities display subtle variations which can be hard for non-native speakers to grasp.

These phenomena, then, illustrate cases where misunderstandings and cross-language confusion are likely to arise even for reasonably competent language learners. To verify that intermediate learners do indeed have problems in these areas, we carried out two small scale investigations.

Corpus analysis: The first move was simply to look at a moderate sized set of essays (13,000 words) produced by Persian-speaking intermediate learners of English and classify the errors. This initial analysis indicated that learners do have errors with these closed-class words, but also showed that Persian learners have an extra problem arising from the very extensive use of 'light' verbs in Persian, i.e. constructions using a common support verb (something like '*make*' or '*do*') and a bare noun where English would use a specific lexical item (Butt 2002).

The analysis of this corpus was fairly informal, and was aimed simply at determining whether the area of closed class words and light verbs was worth further investigation. In particular, we wanted to see whether intermediate learners do have difficulty with choosing appropriate prepositions and tense/aspect markers. Simple inspection suggested they do, but, as noted above, if you do not have direct access to the message that someone is trying to convey then you cannot be absolutely certain that what they have said fails to convey it.

Online translation exercises: To back up these informal observations, we constructed an on-line set of 30 translation exercises and invited Persian-speaking learners of English to attempt them. This was not a very controlled experiment, since the participants were self-selected from the community of Iranian students in Manchester, but it did give us a useful collection of data to work from. 48 people filled in the form, with 34 attempting at least 20 of the exercises and about 6 of the others attempting at least 10 (see Appendix for an example). We got an average of about 20 distinct translations for each of the target sentences, with a maximum of 33 distinct translations for '*Ali zmyt sxty xvrd*' – '*Ali earth hard hit*', or '*Ali fell badly*' and a minimum of 16 for '*aEtqad bh xda darm*' – '*belief in God I-have*' ('*I believe in God*'). Typically a small number of these contained such severe grammatical errors that the parser could not do anything with them even after constraint relaxation (not more than two or three in each case). Of the others, a very small number contained gross lexical errors (e.g. '*he rode the bus*'

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

for a sentence whose correct translation was '*he got into the car*'). About half the remainder were acceptable translations, and the rest displayed the kinds of error in closed class words ('*he got onto the car*' rather than '*he got into the car*', '*he is shouting at everybody*' rather than '*he shouts at everybody*'...) and referential form (notably omission of '*the*') that we were expecting.

These experiments were not intended as a definitive investigation into the nature of the errors produced by Persian speaking learners of English, but they do confirm that there are students who have a reasonable level of English but who produce errors that arise from the partiality of the mappings between closed class items in the two languages. The fact that someone is not entirely confident about which preposition or which form of the verb to use does not, in general, prevent them from producing something meaningful in quite the way that a lack of an open class word does. But it does mean that they sometimes produce utterances which convey something like what they meant to say rather than exactly what they meant, and a tool which can help them see cases where they have made errors of this kind could well prove useful.

1.2 How can you tell whether what someone says is what they meant?

When you are trying to provide a tool to help people learn grammar, there is a (fairly) clear test: is what they have produced grammatical? If a learner says '*I is reading a book about language teaching*', you can be reasonably sure that they have made a grammatical mistake, and that they meant to say '*I am reading a book about language teaching*'[2] But if they do say '*I am reading a book about language teaching*,' you simply have to accept that there is nothing wrong with what they have said.

In general, you cannot tell whether what someone says is what they meant to say unless you know what they meant to say. In some cases it can be hard to construct a reasonable interpretation, but given an appropriate context almost anything can be made to sound reasonable (or, at worst, metaphorical).

It is, therefore, fairly hard to see how you could write a program which looked, in isolation, at what someone said and reported on how successfully it expressed their intended message. In order to know whether '*I live in Derbyshire*' or '*I am living in Derbyshire*' more accurately reports my domestic arrangements, you have to know whether I expect to move somewhere else in the near future. In order to know which of '*I am reading a book about language learning*' and '*I am writing a book about language learning*' is a better description of what I am doing you need to know what I am doing. *In order to know whether my words express my thoughts, you have to know what my thoughts are*.

It is, of course, extremely difficult to know what someone else is thinking. What else is language for, other than to try to reveal your thoughts to someone else? So how can we get a program to ascertain what someone is thinking, and hence to use that as a basis for working out whether what they have said is what they meant to say?

One way of approaching the problem is to set translation exercises. If I say '*Translate the phrase 'I hate you' into French*' then I have a reasonably reliable handle on what

[2]or possibly, but less likely, 'He is reading a book about language teaching.'

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

ideas you are trying to express, and I can use this to decide whether '*Je t'adore*' expresses your thoughts accurately or not. The framework of the work reported here, then, is as follows:

1. The teacher provides a set of translation exercises. These are given as sentences in the learner's native language (L1), so that we can be fairly sure that there will be no misunderstandings at this point: the idea conveyed by the L1 sentence is the idea the learner wishes to express. The teacher does not have to supply candidate translations.
2. The learner provides a putative translation into the target language (L2).
3. The computer constructs meaning representations for the original L1 sentence and the L2 translation, as described in Section 2.
4. The two meaning representations are compared, using rules which describe the relationship between complex combinations of L1 and L2 terms (Section 3), and information obtained from this comparison is given back to the learner (Section 4).

The architecture of the system, then, is as shown in Figure 1. A tool of this kind has a number of attractive features:

1. By concentrating on translation exercises we can be fairly sure that we know what the student is trying to say. This is important: if we do not know what someone is trying to say, how can we determine whether they are saying it correctly? The point is not that advanced students generally produce, or should produce, utterances in the target language by translating them from the equivalent in their native language. Rather, we are using this form of exercise as a way of getting a handle on what the student is trying to say. If the task is to produce a translation then the content of the translated text should be the same as the content of the source. Since students are translating from their native language into the target language they should not have any difficulty in understanding the source text. So any discrepancies that arise between the source and target can reasonably be assumed to arise from lack of understanding of the lexical, sub-lexical and structural choices made in constructing the target.
2. Students seldom produce translations which are syntactically more complex than

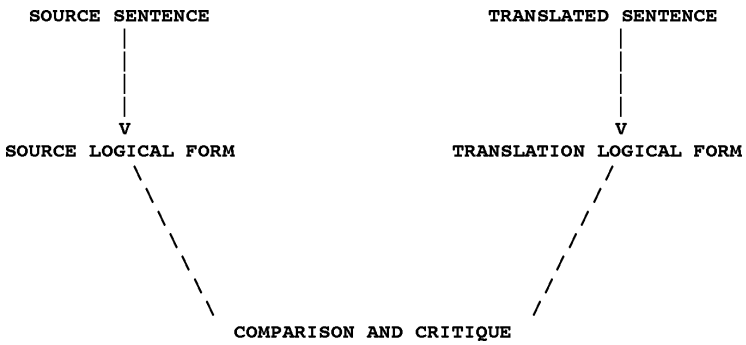


Fig. 1. Architecture of the tool.

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

the source material. We can thus control the syntactic complexity of the student input, which helps keep the processing time and the number of potential analyses manageable. This is particularly significant in the context of learner input. The students we are targetting do still make occasional grammatical errors. We therefore need to allow for a certain amount of soft parsing, which inevitably increases the number of potential interpretations. If we can keep the target output simple then the number of spurious soft analyses will be kept to a minimum.

3. The teacher does not need to provide translations for the source texts. We do need an underlying set of bilingual equivalences, but once we have this the translations of individual sentences can be checked automatically. This saves the teacher a considerable amount of effort, and indeed opens up the possibility of getting students to set their own exercises or to set exercises for one another without any intervention from the teacher.

2 Logical forms

The first task, then, is to construct meaning representations of the two sentences. Exactly what should go into a meaning representation is open to debate. We follow fairly widespread practice in constructing 'logical forms' (LFs) – paraphrases in some appropriately chosen formal language which have the same consequences as the original sentence.

There is almost nothing else you can do. Any computational representation is, by definition, an expression in a formal language. You can make good or bad choices of formal language, and you can capture more or less of the content of the original sentence in your paraphrase, but you cannot make a computational meaning representation which is not an expression in some formal language.

The specific formal language we choose for this task is a dynamic form of Turner's (1987) 'property theory'. The precise reasons for using this language are discussed in Ramsay and Seville (2000). The analysis in Figure 2 for '*She shouts at everyone*' illustrates a number of the key points. The key points about this LF are as follows:

- referring expressions are denoted by 'reference terms', such as the term `ref(lambda(E, centred(E) & f(E)))`, which corresponds to the pronoun '*she*'. This term can be seen as an instruction to find an item in the discourse which satisfies the property of being both 'centred' and female. Finding a suitable item is known as 'anchoring' it (Barwise & Perry, 1983).
- the temporal properties of the event are specified by introducing a relation

```
forall(A :: {one(A)}
    exists(B :: {B is interval
                 & ends_after(ref(lambda(C, now(C))), B)}
           exists(D :: {aspect(simple, B, D)}
                    theta(D,agent,ref(lambda(E,centred(E)&f(E))))
                    & shout(D) & D is event & at(D, A)
```

Fig. 2. Logical form for 'she shouts at everyone'.

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

between speech time, denoted by the reference term `ref(lambda(C,now(C)))`, and 'reference time' `B`, and between `B` and the event `D` itself.

- The relations between the event and the other items are specified either using 'thematic roles', such as the relation given in `theta(D, agent, ref(lambda(E, centred(E) & f(E))))`, which says that '*she*' is the agent of the event, or by using simple relations such as `at(D, A)`. The number and nature of the thematic roles that are deployed in a given language is subject to considerable debate (Dowty, 1989, 1991; Pollard & Sag, 1987; Wechsler, 1995). We take a fairly minimalist view, informed by a mixture of syntactic facts and semantic intuitions.

These LFs are constructed using well-known 'compositional' techniques, where meanings of complex utterances are built up by systematic combination of meanings of simpler parts. The specific approach we use employs ideas similar to the 'glue languages' of Dalrymple *et al*. (1996) and van Genabith and Crouch (1997), but the important point to note is that the approach is appropriate for any language for which we have a syntactic analysis. So we can build LFs for Persian sentences just as easily as for English ones, so long as we have a parser for Persian.

The grammar we use for Persian is described in Mirzaeian (2003). The details are outside the scope of the current paper. All that matters here is that we do have a parser for Persian, which means that we can construct LFs for Persian sentences, as seen in Figure 3.

(1) *av sr hmh dad myznd*. (she/he at everyone shout hits)

Note that there are no elements in the LF directly corresponding to the noun '*dad*' and the verb '*myznd*'. Instead we have extracted the compound item '*zn-dad*', made up of the stem of '*myznd*' ('*my-*' is the present tense marker) and the noun '*dad*', to show that what we have here is a light verb denoting a specific kind of event.

3 Comparing logical forms

The LFs in Section 2 contain a great deal of fairly subtle information – descriptions of the temporal structure of events, relations between entities, referential and quantificational information. Just as with the original natural language utterance, some of this is encoded in the structure of the representation. Relative scopes of logical operators, for

```
exists(A :: {A is interval & ends_after(ref(lambda(B, now(B))), A)}
       forall(C :: {one(C)}
              exists(D :: {aspect(simple, A, D)}
                     theta(D,
                           agent,
                           ref(lambda(E, centred(E) & thing(E))))
                     & zn-dad(D) & D is event & sr(D, C)
```

Fig. 3. Logical form for (1).

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

instance, are indicated by their positions within the LF, as shown in Figure 4, and information about salience is encoded in the left to right order of its constituents, which reflects the left to right order of the original utterance.

If we want to extract every nuance of the meaning of the utterance, complex structured representations like these are essential – if, for instance, you do not get the relative scope of '*not*', '*any*' and '*every*' right then you have no chance of telling the difference between the two sentences in Figure 4. But if we want to compare two putatively equivalent LFs, the structure of the representation can get in the way. If we want to compare LFs it makes life easier if we can convert them to some more uniform representation, whilst losing as little as possible of the content captured by the originals.

We therefore apply a number of standard logical transformations to convert the original interpretation to a format known as 'clausal form'. The result is a very flat representation, usually just a list of individual facts, which has a great deal in common with the representations obtained in 'minimal recursion semantics' (Copestake *et al.*, 1998; Copestake, Flickinger & Sag, 1999). Figure 5 shows the flattened form of the LF for '*she shouts at everyone*'.

This flat logical form (FLF) starts with a number of simple facts, together with the statement `OLD(#4)`, which says that the object whose internal name is `#4` is already present in the context. In other words, something will be OLD if it corresponds to a referring expression in the text.

The last part of the representation in Figure 5 is a rule which says that whatever 'one' you choose, there is a shouting event with some known feminine entity `(f(#6), OLD(#6))` as the agent and your chosen one as the thing being shouted at. This is just

```
not(forall(A :: {one(A)}
    exists(B :: {aspect(simple, ref(lambda(C, now(C))), B)}
                        theta(B,
                              agent,
                              ref(lambda(E, centred(E) & speaker(E))))
                        & hate(B) & B is event & theta(B, object, A))

forall(A :: {one(A)}
   not(exists(B :: {aspect(simple, ref(lambda(C, now(C))), B)}
                           theta(B,
                                 agent,
                                 ref(lambda(E, centred(E) & speaker(E))))
                           & hate(B) & B is event & theta(B, object, A))
```

Fig. 4. 'I don't hate everyone' vs 'I don't hate anyone'.

```
[#2 is interval, now(#4), ends_after(#4, #2), OLD(#4),
  [CHOOSE(#5), one(#5)]
  => [aspect(simple, #2, #3(#5)), centred(#6), f(#6), OLD(#6),
       theta(#3(#5),agent,#6), shout(#3(#5)),#3(#5) is event, at(#3 (#5),#5)]
```

Fig. 5. Clausal form for 'she shouts at everyone'.

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

```
[#0 is interval,now(#2),ends_before(#0,#2),aspect(perfect,#0,#1),
  centred(#3), thing(#3), theta(#1, topic, #3), predication(#1),
  #1 is event, maSyn(#4), svar(#1, #4), OLD(#4), OLD(#3), OLD(#2)]

[#6 is interval,now(#8), ends_before(#6,#8), aspect(simple, #6, #7),
  centred(#9),m(#9),theta(#7,topic,#9), predication(#7), #7 is event,
  car(#10), into(#7, #10), OLD(#10), OLD(#9), OLD(#8)]
```

Fig. 6. 'svar maSyn Sd' vs 'he got into the car'.

about equivalent to the original analysis in Figure 2, but is clearly easier to compare with a similar representation of a Persian sentence.

3.1 Simple cases

Analysing whether an English sentence input by a learner is a reasonable translation of a Persian source sentence thus becomes a matter of comparing lists of the kind given in Figure 5. We will illustrate how we do this with a very simple pair:

(2) (a) *svar maSyn Sd*.
(b) *he got into the car*.

The FLFs for these are shown in Figure 6. These two are close to word-for-word translations of each other. The first few elements, which correspond to the tense markers, are in fact identical, apart from the name changes for the various entities concerned. We then get items which encode the aspect markers, which are different because the two languages have slightly different aspectual systems. The Persian has `centred(#3), thing(#3)` for the subject, where the English has `centred(#9), m(#9)`: the difference arises becauses the Persian has no explicit subject, and the verbal agreement marking tells us nothing about gender, so we cannot tell whether it should have been '*he*', '*she*' or '*it*', whereas the English pronoun '*he*' is clearly marked as masculine.[3] The next few items again match directly, since both sentences are essentially copula sentences, so the items involved are in the same relationships. We then get a couple of items where the English is a translation of the Persian ('*maSyn*' = '*car*', '*svar*' = '*into*') and then a set of statements about which items are already in the context.

So to verify that (2b) is a reasonable translation of (2a) we just need to know that the Persian past perfect can be translated using the English simple past; that the Persian zero subject denotes something of unknown gender, so that an entity marked in Persian as being `thing` and `centred` may correspond to an entity which is marked in English as `m` and centred (note that this only applies if both are marked as centred); and that '*car*' is a legitimate translation of '*maSyn*', and likewise for '*into*' and '*svar*'.

This is a particularly simple example, but it nonetheless illustrates some important points:

- By looking at semantic representations, we are not fooled by the different surface

[3]centred(X) is used for items which have been referred to by pronouns or similar constructions, and which hence should be particularly salient in the context (Ramsay & Seville 2000).

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

```
[maSyn(X)] && [car(Y)] : CONTEXT :-
       matchArgsInContext([X], [Y], CONTEXT).
```

Fig. 7. 'maSyn'='car'.

structures of the two languages.

- We can isolate the specific components of a translation which are problematic. If the English translation had been '*he was getting into the car*' the aspectual mismatch would have shown up straight away, without obscuring the fact that the remainder of the translation was fine.
- We need sets of bilingual equivalents.

Simple bilingual equivalents are given by rules like Figure 7. The first part of this rule says that if the Persian FLF contains a set of items that match the left hand side and the English one contains elements that match the right hand side then the English is potentially an acceptable translation of the Persian – that if the Persian contains the term '*maSyn*' and the English contains '*car*' then these two items can be matched. The remainder is there to ensure that we keep track of which items in the two FLFs have been matched – crucially in this case that the maSyn and the car are the same thing.

The result of comparing these two interpretations, then is a triple `{P, E, C}` where `P` is a set of terms that appear in the Persian interpretation but which have no equivalents in the English, `E` is a set of terms that appear in the English but have no equivalents in the Persian, and `C` is a set of 'comments'. These comments are added by rules which deal with partial matches, to provide a record of places where a mismatch between the Persian and the English interpretations has arisen from some common misunderstanding. Given '*svar maSyn Sd.*' and '*He got into the car*', for instance, we get `{[], [], [perfect < simple]}`, indicating that the English contains an appropriate translation of the Persian except for the comment that the English simple past is not always a perfect match for the Persian past perfect (in some cases the English past perfect is to be preferred, but it is very difficult to characterise the situations where this holds).

The process of comparing logical forms depends crucially on the presence of such equivalence rules. If we do not tell the system what the English equivalent of '*maSyn*' is then it will not have much chance of telling whether what the student has written is a reasonable translation! Providing a substantial bilingual dictionary is a major undertaking, but we do not need a full set of Persian-English equivalents. The tool described here is used for assessing putative translations of exercises which have been set by the teacher as part of a general course. As such, it is not unreasonable for the teacher to provide the English equivalents *for the words that appear in the exercises*, at least in simple cases where there is a word-for-word equivalence. Over a period of time, such a dictionary may grow quite large, but at any one moment it only needs the vocabulary required for the currrent exercise, and this can be added incrementally without undue effort.

3.2 Near misses

In many cases, however, the student's translation will contain a word which is in some way related to what was expected but is not an exact match. There are a number of possibilities:

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

- the two words may have much the same meaning but be in different registers (e.g. slang).
- the term in the intended translation may be a superordinate or subordinate of the term in the source ('*vehicle*' instead of '*car*', for instance).
- the term in the translation may mean the same as the one in the source in some contexts, but not in the context of the current sentence.

We deal with the first two of these by maintaining *monolingual* sets of lexical relations for the target language, where a relation specifies the two words and a label which explains how they are related. So for the word '*car*', for instance, we might record the following relations: `rel(car, motor, slang), rel(car, vehicle, super), rel(hatchback, car, super) ...` (note that we can perfectly easily infer that '*hatchback*' is a subset of '*car*' from the fact that 'car' is a superset of '*hatchback*', so we do not need to record hyponyms as well as hypernyms). Then in situations where we have some lexical item in the translation which does not match anything in the source, we can look for entities which are related to it which *do* correspond to items which occur in the source.

We only need such a set for the target language. Errors of this kind tend to indicate lexical gaps in the learner's knowledge, and learners do not generally have such gaps in their native language – Persian speakers will know what '*maSyn*' means. What they may not know is that its English translation is '*car*', in which case they may use an English word which they do know is the translation of some related Persian term. So all we need is the set of relations for English, and we can use WordNet (Fellbaum, 1998) for at least the superset relations for English.

We further allow the transitive closure of the set of lexical relations, so that if someone translates '*maSyn*' as '*thing*' we can report that they have used a superordinate of something which is itself a superordinate of the target term. So if the English translation had been '*He got into the vehicle*', the resulting triple would have been `{[], [], [perfect<simple, [rel(vehicle, car, super)]]}`, showing that it is not too bad but that a superordinate of the target word has been used; and if the proposed translation had been '*He got into the thing*' then the triple would have been `{[], [], [perfect<simple, [rel(thing, vehicle, super), rel(vehicle,car,super)]}`, showing that the word '*thing*' is two steps removed from the correct term (note that we have to be careful when using the WordNet hypernyms, since WordNet makes use of intermediate concepts without direct lexical realisations. So if someone writes '*vehicle*' where the correct translation is '*car*', we will obtain the chain `[car,motor_vehicle,wheeled_vehicle,vehicle]` from WordNet. We omit the hyphenated elements of such a chain from our diagnostic reports.

3.3 Phrasal equivalents

The fact that the semantic analysis abstracts from the surface form is particularly important. Consider (3) and (4):

(3) *aEtqad bh xda darm.*
(4) *I believe in God.*

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

The Persian here is 'belief in God have-I' – very similar to 'I have belief in God', but quite different from the surface form 'I believe in God'. The FLFs for these two are shown in Figure 8.

The material relating to tense and aspect can clearly just be ticked off point by point – both sentences are just simple present tense, and there is no reason not to match these parts of the FLF. Likewise, the role of the speaker in each case is the same and can be dealt with very straightforwardly.

The first complication arises from the fact that the Persian reports a '*having*' event `dar/St(#16)` whose `object #20`[4] is a belief in God, whereas the English has a '*believing*' event which is in God.

To cope with this, we need rules which match several elements of the two FLFs at a time – we need to be able to say that '*having a belief in X*' is the same as '*believing in X*'. So we need rules like the one in Figure 9.

The first part of this rule provides an equivalence between sets of terms in the two logical forms. Apart from the fact that there are several items on each side of the equivalence, this is very similar to the rule in Figure 7, but it is clearly even more important to match up the arguments appropriately when we are matching several terms from each side.

The second problem with Figure 9 arises from the form of the word '*aEtqad*'. There is no difference in written Persian between definite and generic noun phrases (NPs), or indeed between definite NPs and nouns. It is therefore impossible to tell whether '*aEtqad*' is a definite NP ('*the belief*'), or a generic NP ('*belief*') or a simple noun marking a light use of the verb '*dar*'. Rather than produce multiple interpretations, alternating between definite and generic readings, we always interpret simple unmarked NPs as being definite, but we then allow for the presence of an unwanted definiteness marker in

```
[#15 is interval, now(#19), ends_after(#19,#15),
  aspect(simple,#15,#16),aEtqad(#20),theta(#16,ra,#20),da{r/St}(#16),
  #16 is event, xda(#21), svar(#20,#21), centred(#22), speaker(#22),
  theta(#16,agent,#22), OLD(#22), OLD(#21), OLD(#20), OLD(#19)]

[#2 is interval, now(#4), ends_after(#4,#2), aspect(simple, #2, #3),
  centred(#5), speaker(#5), theta(#3,agent,#5), believe(#3),
  #3 is event, named(#6, God), in(#3, #6), OLD(#6), OLD(#5), OLD(#4)]
```

Fig. 8. 'aEtqad bh xda darm.' vs 'I believe in God.'

```
[aEtqad(X), bh(X, Y), theta(Z,ra,X), dar(Z)]
~~
[believe(A), in(A,B)]
: CONTEXT :-
matchArgsInContext([Z, Y], [A, B], CONTEXT).
```

Fig. 9. 'having belief in' = 'believing in'.

[4]The postposition 'ra' marks the direct object in Persian.

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

```
[aEtqad(X), bh(X, Y), theta(Z,ra,X), dar(Z), OLD(X)]
~~
[believe(A), in(A,B)]
: CONTEXT :-
matchArgsInContext([Z, Y], [A, B], CONTEXT).
```

Fig. 10. 'having the belief in' = 'believing in'.

cases like Figure 8 by adapting the corresponding rule, so that Figure 9 is actually as in Figure 10.

3.4 Light verbs

The widespread use of light verbs in Persian is particularly problematic. Persian makes extensive use of constructions made up of a fairly general verb (something like '*make*', '*do*', '*hit*'...) and an uninflected noun where English has a single lexical item. The English verb '*shout*', for instance, corresponds to the combination '*dad myznd*' ('*shout hit*'), the verb '*email*' corresponds to '*aymyl myznd*' ('*email hit*'), and so on. These are problematic for learners, because it is hard to know whether a literal translation will suffice (because the corresponding English phrase is acceptable, e.g. '*shall we do lunch?*') or whether there is a preferred lexical verb. And they are problematic for automatic diagnosis because a simple-minded translation will leave us with an NP which is problematic for two reasons: (a) it will look like a definite NP, since the default in Persian is that unmarked NPs are definite, but it cannot be successfully dereferenced, and (b) it makes no real contribution to the interpretation.

We deal with this by allowing the form of the NP to select different senses of the verb, so that the phrase '*dad myznd*' is seen as reporting an event of type *zn-dad*, whereas '*aymyl myznd*' reports an event of type *zn-aymyl*. At the same time, we eliminate the contribution that the NP makes to the overall interpretation, so that the interpretation of '*av sr hmh dad myznd*.' is as given in Figure 3, Section 2, with the compound of '*dad*' and '*myznd*' giving rise to a single item which can be straightforwardly matched to the verb '*shout*' (and thence, by English lexical relations, to other words like '*yell*' and '*scream*').

4 Conclusions and further work

The mechanisms outlined in Section 2 provide diagnostic information about mismatches between the user's intended English translation and the original Persian sentence. This information comes in the form of a list of unmatched terms from the logical forms, possibly supplemented by comments supplied by specific rules. This is some way from what a learner wants to see, just as the output of a soft parsing algorithm that records a list of violated constraints is unlikely to be immediately helpful to the average learner. We therefore need to find some way of turning this rather terse information into meaningful messages.

Exactly what these meaningful messages should look like is still unclear. It is easy enough to generate a readable string such as '*English progressive used for Persian*

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

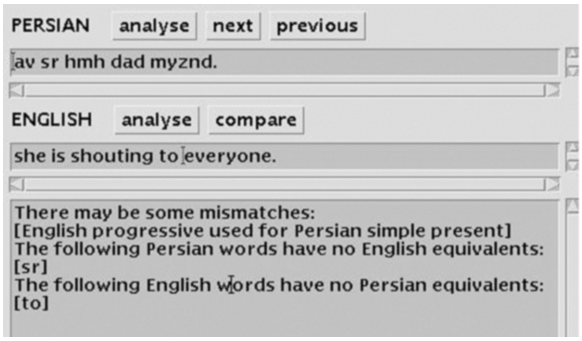


Fig.11. Screenshot of diagnostic messages.

simple present' from a cryptic code such as `[!(prog=simple)]`, as shown in Figure 11, but we need to carry out further investigations into what learners want to be told, and when they want to be told it, before we can be sure about the best form of feedback to provide.

The tool described above, then, shows how we can generate diagnostic messages based on matching the logical forms for a target sentence and for a putative translation. Using translation exercises for this task serves a number of purposes:

- It gives us a window onto the learner's intended message. This is crucial for any attempt to carry out semantic diagnosis: we cannot tell whether what a learner produces is an accurate realisation of what they want to say unless we know what they want to say. If they are trying to carry out a translation exercise then we can be fairly confident that the content of the source text is indeed the message that they are trying to convey.
- Letting the teacher set the exercises means that they can focus on issues that are currently being taught. But since the diagnosis is carried out automatically, the teacher does not need to provide sample correct translations, and hence there is scope for students to construct a range of possible solutions, not all of which may have been anticipated by the teacher. There is still, of course, a need for a set of bilingual equivalents, but because these are specified for individual words or phrases they do not need to be revised every time a new exercise is set.
- Letting the teacher set the exercises also provides control over the likely complexity of the students' solutions, since the translation of a single sentence is generally going to be of similar length and complexity to the target. This helps improve the performance of the tool, since short sentences are generally analysed considerably more quickly, and with considerably fewer potential analyses, than long ones.

It is worth noting at this point that concentrating on finding mismatches between pairs of supposedly equivalent sentences helps us to sidestep the thorny problem of ambiguity resolution. It is generally the case that wide-coverage Natural Language Processing (NLP) systems find numerous unexpected analyses of quite simple sentences, and there

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

are no generally reliable mechanisms for choosing the correct interpretation. In the present context, simply finding the analyses with the fewest mismatches is a fairly sensible strategy. If the source sentence has S1, S2, ..., Sn as potential interpretations and the target has T1, T2, ..., Tm, where Si and Tj have no mismatches, then it is reasonable enough to suppose that the student took Si as the interpretation of the source and intended Tj as the interpretation of the translation. This means that as far as this task is concerned we can forget about trying to find the 'correct' interpretations of the source and its translation and just look for the ones with the fewest mismatches, which is a much easier task.

The resulting system provides a useful starting point for investigations into the automated diagnosis of learners' semantic errors. We need to carry out further investigations into the use of such a tool in the classroom, but the engine described above does seem to have a number of innovative applications.

References

Barwise, J. and Perry, J. (1983) *Situations and Attitudes*. Cambridge, MA.: Bradford Books.

Butt, M. (2002) The light verb jungle. In: *Workshop on light verbs*, Harvard.

Copestake, A., Flickinger, D., Malouf, R., Riehemann, S. and Sag, I. (1998) Translation using minimal recursion semantics. *Proceedings of the Sixth International Conference on Theoretical and Methodological Issues in Machine Translation*, Leuven.

Copestake, A., Flickinger, D. and Sag, I. (1999) Minimal recursion semantics: an introduction. *Technical report*, Stanford University.

Dalrymple, M., Lamping, J., Pereira, F.C.N. and Saraswat,V. (1996) A deductive account of quantification in LFG. In: Kanazawa, M., Piñón,C. and de Swart, H. (eds.), *Quantifiers, Deduction and Context*. Stanford: CSLI publications, 33–58.

Dowty, D. R. (1989) On the semantic content of the notion of 'thematic role'. In: Chierchia, G., Partee, B. H. and Turner, R. (eds.), *Properties, Types and Meaning II: Semantic Issues*. Dordrecht: Kluwer Academic, 69–130.

Dowty, D. R. (1991) Thematic proto-roles and argument selection. *Language*, **67**: 547–619.

Fellbaum, C. (ed.) (1998) *WordNet: An electronic lexical database*. Cambridge, MA: Bradford Books.

Flickinger, D., Oepen, S., Uszkoreit, H. and Tsujii, J. (2000) Special issue on efficient processing with HPSG: Methods, systems, evaluation. *Journal of Natural Language Engineering*, **6**(1).

Mirzaiean, V. (2003) Content-based support for Persian learners of English. PhD thesis, UMIST.

Pollard, C. J. and Sag, I. A. (1987) An Information Based Approach to Syntax and Semantics: Vol 1 Fundamentals. *CSLI lecture notes 13*, Chicago: Chicago University Press.

Ramsay, A.M. and Seville, H. (2000) Models and discourse models. *Journal of Language and Computation*, **1**(2): 167–181.

Turner, R. (1987) A theory of properties. *Journal of Symbolic Logic*, **52**(2): 455–472.

van Genabith, J. and Crouch, R. (1997) How to glue a donkey to an f-structure. In: Bunt, H., Kievit, L., Muskens, R. and Verlinden, M. (eds.), *2nd International Workshop on Computational Semantics*. University of Tilburg, 52–65.

Appendix: A sample translation set

01 = he entered to the car
02 = I belive in God.

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press

03 =
04 = email me please
05 = he made a table by wood.
06 = Ali fell down very hard
07 = he has cancer
08 = at last they did
09 = he racked in the jail
10 = he brings the chair into the house
11 = she becomes a good secretery
12 = he bit me
13 = Hassan bit me at the golf game
14 = this dress suits for you
15 = Mohamad for 5 min. belived I am right, but I was not.
16 = the car damaged very bad on the casulty
17 = I sved her husband from troble
18 = when do the boxes will be recived?
19 = the students under took the exam yesterday
20 = I retearned the suit which I bought
21 = Kimia is listening to the radio
22 = Teacher got the exam from students
23 = our visa will be expired september
24 = Ali always makes joke with Hosein
25 = he missed his parents when he was child
26 = Nader was sweeping when I entered
27 = Guilty accepted his crim in the police presence
28 = some soldger racked Hassan in the jail
29 = he taught every one talk to him in order to pay less
30 = British asked Russia's cooperation in midle east crises

https://doi.org/10.1017/S0958344005001011 Published online by Cambridge University Press