Towards AMR-BR: A SemBank for Brazilian Portuguese Language

Rafael Torres Anchiêta, Thiago Alexandre Salgueiro Pardo
Interinstitutional Center for Computational Linguistics (NILC)
Institute of Mathematical and Computer Sciences, University of São Paulo
rta@usp.br, taspardo@icmc.usp.br

Abstract

We present in this paper an effort to build an AMR (Abstract Meaning Representation) annotated corpus (a semantic bank) for Brazilian Portuguese. AMR is a recent and prominent meaning representation with good acceptance and several applications in the Natural Language Processing area. Following what has been done for other languages, and using an alignment-based approach for annotation, we annotated the Little Prince book, which went into the public domain and explored some language-specific annotation issues.

Keywords: Abstract Meaning Representation (AMR), corpus annotation, Portuguese language

1. Introduction

Due to its wide applicability and potentialities, Natural Language Understanding (NLU) has gained interest and fostered research on themes of computational semantics (Oepen et al., 2016). According to Ovchinnikova (2012), NLU is the field of Natural Language Processing (NLP) that deals with machine reading comprehension. The objective of an NLU system is to specify a computational model to interpret one or more input text fragments. The interpretation is usually carried out by a semantic parsing technique, which maps natural language into a suitable meaning representation.

A meaning representation is one of the most important components in semantic parsing. Its production is motivated by the hypothesis that semantics may be used to improve many natural language tasks, such as summarization, question answering, textual entailment, and machine translation, among others. In this context, there are several available meaning representations, as the traditional First-Order Logic (FOL), as detailed in Jurafsky and Martin (2009), semantic networks (Lehmann, 1992), Universal Networking Language (UNL) (Uchida et al., 1996), and, more recently, the Abstract Meaning Representation (AMR) (Banarescu et al., 2013).

In particular, AMR got the attention of the scientific community due to its relatively simpler structure, establishing the connections/relations among nodes/concepts, making them easy to read. Moreover, AMR structures are arguably easier to produce than traditional formal meaning representations (Bos, 2016).

According to Banarescu et al. (2013), AMR-annotated corpora are motivated by the need of providing to the NLP community datasets with embedded annotations related to the traditional tasks of NLP, for instance, named entity recognition, semantic role labeling, word sense disambiguation, and coreference. In this sense, the AMR annotation especially focuses on the predicate-argument structure as defined in PropBank (Kingsbury and Palmer, 2002; Palmer et al., 2005). Another characteristic of AMR annotation is that words that do not significantly contribute to the meaning of a sentence (which are referred as "syntactic sugar" in the original paper) are left out of the annotation, as articles and the infinitive particle "to".

From the currently available datasets, many semantic parsers emerged (Flanigan et al., 2014; Wang et al., 2015; Peng et al., 2015; Goodman et al., 2016; Zhou et al., 2016; Damonte et al., 2017). Furthermore, with the available parsers, some applications were developed for summarization (Liu et al., 2015) and text generation (Pourdamghani et al., 2016; Song et al., 2017), entity linking (Pan et al., 2015; Burns et al., 2016), and question answering (Mitra and Baral, 2016), among others.

Although there are some available annotated corpora, most of them are for English, producing a gap between English and other languages. In addition, creating such corpora is a very expensive task. For instance, Banarescu et al. (2013) took from 7 to 10 minutes to annotate a sentence in AMR representation. However, in spite of the difficulties, it is important to put some effort on corpus creation for other languages. Annotated corpora are important resources, as they provide qualitative and reusable data for building or improving existing parsers, and for serving as benchmarks to compare different approaches.

In order to fulfill this gap, we annotated a corpus in AMR representation for the (Brazilian) Portuguese language, which we report in this paper. In addition, we also detail some differences between Portuguese and English AMR annotations. To the best of our knowledge, this is the first initiative on AMR for Portuguese. We believe that the availability of such a semantic bank[1] in Portuguese will result in new semantic parsers for this language and support the development of more effective NLP applications.

In the following section, we briefly introduce the AMR fundamentals. In Sections 3 and 4, we present our corpus and report the annotation process and its results. Section 5 concludes the paper.

2. Abstract Meaning Representation

Abstract Meaning Representation (AMR) is a semantic representation language designed to capture the meaning of a sentence, abstracting away from elements of the surface syntactic structure, such as part of speech tags, word ordering, and morphosyntactic markers (Banarescu et al., 2013). It may be represented as a single-rooted acyclic directed

[1] A "SemBank", as referred in one of the first AMR papers.

graph with labeled nodes (concepts) and edges (relations) among them in a sentence. AMR concepts are either words (e.g., "girl"), PropBank framesets ("adjust-01"), or special keywords such as "date-entity", "distance-quantity", and "and", among others. PropBank framesets are essentially verbs linked to lists of possible arguments and their semantic roles. In Figure 1, we show a PropBank frameset example. The frameset "**edge.01**", which represents the "move slightly" sense, has six arguments (Arg 0 to 5).

Frameset **edge.01** "move slightly"

Arg0: causer of motion	Arg3: start point
Arg1: thing in motion	Arg4: end point
Arg2: distance moved	Arg5: direction

Ex: [$_{\text{Arg0}}$Revenue] *edge* [$_{\text{Arg5}}$ up] [$_{\text{Arg2-EXT}}$ 3.4%] [$_{\text{Arg4}}$ to $904 million] [$_{\text{Arg3}}$ from $874 million] [$_{\text{ArgM-TMP}}$ in last year's third quarter]. (wsj_1210)

Figure 1: A PropBank frameset (Palmer et al., 2005)

For semantic relationships, besides the PropBank semantic roles, AMR adopts approximately 100 additional relations, as general relations (e.g., :mod, :location, :condition, :name, and :polarity), relations for quantities (:quant, :unit, and :scale) and for dates (:day, :month, and :year), among others.

AMR may also be represented in two other notations: in first-order logic or in the PENMAN notation (Matthiessen and Bateman, 1991). For example, Figures 2 and 3 present the canonical form in PENMAN and graph notations, respectively, for the sentences with similar senses in Table 1.

Sentences
The girl made adjustment to the machine.
The girl adjusted the machine.
The machine was adjusted by the girl.

Table 1: Sentences with the same meaning

```
(a / adjust-01
      :ARG0 (g / girl)
      :ARG1 (m / machine))
```

Figure 2: PENMAN notation

AMR assigns the same representation to sentences that have the same basic meaning. Furthermore, as we may observe in the example, the concepts are "adjust-01", "girl", and "machine", and the relations are :ARG0 and :ARG1, represented by labeled directed edges in the graph. In Figure 2, the symbols "a", "g", and "m" are variables, which may be re-used in the annotation, corresponding to reentrancies (multiple incoming edges) in the graph.

Moreover, AMR represents negation in a different way. It uses the *:polarity* relation between the negated concept and the constant '−' (minus signal). For instance, the sentence "I do not much like to take the tone of a moralist.", extracted

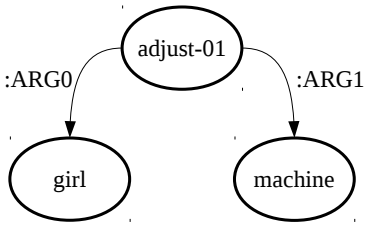

Figure 3: Graph notation

from the Little Prince book, produces the PENMAN notation in Figure 4.

```
(l / like-01 :polarity -
      :ARG0 (i / i)
      :ARG1 (t / take-01
            :ARG0 i
            :ARG1 (t1 / tone
                  :poss (m / moralist)))
      :degree (m1 / much))
```

Figure 4: PENMAN notation representing negation

Finally, to evaluate the AMR structures, Cai and Knight (2013) introduced the Smacth metric, which computes the degree of overlap between two AMR structures, computing precision, recall, and f-score over AMR annotation triples.

3. Our Corpus

There are some available corpora in the Linguistic Data Consortium (LDC), which offer texts in different domains but are not freely available. For now, only two AMR corpora are publicly accessible[2]: Bio AMR Corpus and the Little Prince Corpus. The first includes texts from the biomedical domain, extracted from PubMed[3], whereas the second contains the full text of the famous novel *The Little Prince*, written by Antoine de Saint-Exupéry. The novel was translated into 300 languages and dialects, including Brazilian Portuguese language. Unfortunately, none of the currently available AMR-annotated corpora are for Portuguese.

In this work, following what has been done for other languages, we annotated a public domain version of the Little Prince book written in Portuguese. As a collateral effect of this decision, we may also compare and analyze the annotation of the resulting parallel corpora, composed by the English (source) and Portuguese (target) versions of the book. The original book is organized into twenty-seven chapters. The English version has 1,562 sentences, while the Portuguese one has 1,527. In our annotation process, we aligned all the Portuguese sentences with the English sentences. Furthermore, we calculated some information about the two corpora, such as number of tokens and types, total number of concepts and relations, and maximum and minimum number of concepts and relations found in a sentence, which we show in Table 2.

[2] https://amr.isi.edu/download.html
[3] https://www.ncbi.nlm.nih.gov/pubmed/

Information	English	Portuguese
Number of tokes	16,998	12,703
Number of types	15,829	12,224
Number of concepts	10,528	7,569
Number of relations	10,245	6,676
Average number of tokens	10.88	8.31
Average number of nodes	6	4
Average number of relations	6	4
Maximum number of concepts	37	21
Minimum number of concepts	1	1
Maximum number of relations	49	25
Minimum number of relations	0	0

Table 2: Information about the corpora

4. The Annotation

As aforementioned, we chose as corpus a public domain version of the Little Prince book written in Brazilian Portuguese. Our corpus annotation strategy basically consisted of "importing" the corresponding AMR annotation for each sentence from the English annotated corpus and reviewing the annotation to adapt it to Portuguese characteristics. Doing this, we expected to save time and effort, as a significant part of AMR annotation is probably language independent. More than this, annotation agreement is minimally guaranteed, as it was already checked for the English annotation.

In this sense, we developed an approach with three steps, using the necessary tools and resources to "connect" the English and Portuguese versions of the corpus. Figure 5 illustrates them.

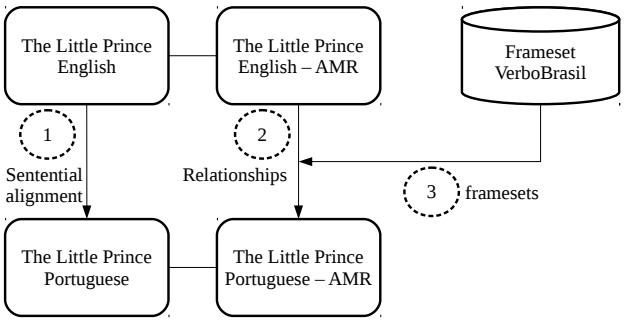

Figure 5: Adaptation of the corpus to the Portuguese language

In the first step, we performed a sentential alignment between the parallel corpora using the TCAlign tool (Caseli and Nunes, 2003), which has a 95% precision. Then, for each sentence, we imported/mapped the AMR relations from the original English sentence to the target Portuguese one. Finally, we included the framesets in each predicate using the *VerboBrasil* dataset (Duran et al., 2013). The *VerboBrasil* dataset is a repository with the sense of verbs in the Portuguese language, similar to the scheme illustrated in Figure 1. This dataset contains examples of a corpus annotated with semantic role labels, created by the PropBank-BR project (Duran and Aluísio, 2012), following the original PropBank initiative. We detail each step in what follows.

Even though the TCAlign tool has 95% precision, we manually checked each alignment, as such information is essential for producing a reliable annotation in Portuguese. We produced 1-1, 1-2, 2-1, 3-1, 1-3, 4-1, 1-4, and 1-5 alignments[4]. As examples, in Tables 3, 4, 5, 6, 7, 8 9, and 10, we present some resulting alignments produced by TCAlign that were manually revised. The overall number for each type of alignment is shown in Table 11. One may also see that there are six sentences in English without correspondence in Portuguese[5].

Source language	Target language
What I need is a sheep.	Preciso é de um carneiro.

Table 3: 1-1 alignment

Source language	Target language
I own three volcanoes, which I clean out every week (for I also clean out the one that is extinct).	Possuo três vulcões que revolvo toda semana.
	Porque revolvo também o que está extinto.

Table 4: 1-2 alignment

Source language	Target language
But I had never drawn a sheep.	Como jamais houvesse desenhado um carneiro, refiz para ele um dos dois únicos desenhos que sabia.
So I drew for him one of the two pictures I had drawn so often.	

Table 5: 2-1 alignment

Source language	Target language
In one of the stars I shall be living.	Quando olhares o céu de noite, porque habitarei uma delas, porque numa delas estarei rindo, então será como se todas as estrelas te rissem!
In one of them I shall be laughing.	
And so it will be as if all the stars were laughing, when you look at the sky at night... you - - only you - - will have stars that can laugh"	

Table 6: 3-1 alignment

In the following steps, we included the sense in each predicate in the sentence, using the *VerboBrasil* dataset, and mapped the relationships to the corresponding AMR relations. Figure 6 shows annotated parallel sentences, in English (left) and in Portuguese (right).

As we see, despite the supposed equality of meaning and annotation, the word '*eu*' (the pronoun "I" in English) does

[4]In an X-Y alignment, X sentences from the original document are aligned to Y sentences in the target one.

[5]Examples of these sentences are "And what good would it do to tell them that?", "Just that.", and "I said.".

Source language	Target language
One sits down on a desert sand dune, sees nothing, hears nothing.	A gente se senta numa duna de areia.
	Não se vê nada.
	Não se escuta nada.

Table 7: 1-3 alignment

Source language	Target language
Hum!.	Hem? respondeu o rei, que consultou inicialmente um grosso calendário.
Hum! ”replied the king; and before saying anything else he consulted a bulky almanac.	
Hum!	
Hum!	

Table 8: 4-1 alignment

Source language	Target language
After that would come the turn of the lamplighters of Russia and the Indies; then those of Africa and Europe, then those of South America; then those of South America; then those of North America.	Vinha a vez dos acendedores de lampiões da Rússia e das Índias.
	Depois os da África e da Europa.
	Depois os da América do Sul.
	Os da América do Norte.

Table 9: 1-4 alignment

Source language	Target language
But in herself alone she is more important than all the hundreds of you other roses: because it is she that I have watered; because it is she that I have put under the glass globe; because it is she that I have sheltered behind the screen; because it is for her that I have killed the caterpillars (except the two or three that we saved to become butterflies); because it is she that I have listened to, when she grumbled, or boasted, or ever sometimes when she said nothing.	Ela sozinha é, porém, mais importante que vós todas, pois foi a ela que eu reguei.
	Foi a ela que pus sob a redoma.
	Foi a ela que abriguei com o pára-vento.
	Foi dela que eu matei as larvas (exceto duas ou três por causa das borboletas).
	Foi a ela que eu escutei queixar-se ou gabar-se, ou mesmo calar-se algumas vezes.

Table 10: 1-5 alignment

not appear in the Portuguese sentence (as it was implicit), but it was annotated. In Portuguese, this phenomenon is called hidden (or implied) subject and it occurs when the subject is not explicit in the sentence but may be easily inferred. In order to keep the similarity with English annotation and the annotation consistency, we annotated all hidden subjects in the Portuguese sentences.

Alignment	Number
1-1	1,356
1-2	41
2-1	60
1-3	3
3-1	10
1-4	1
4-1	1
1-5	1
1-0	6

Table 11: Overall number of alignments

What I need is a sheep	**Preciso é de um carneiro**
(n / need-01	(p / precisar-01
:ARG0 (i / I)	:ARG0 (e / eu)
:ARG1 (s / sheep))	:ARG1 (c / carneiro))

Figure 6: Annotation of parallel sentences

In addition to the subject omission, there are some other differences in the translation into Portuguese. Consequently, the annotation for Portuguese sometimes becomes different from English. In some cases, translations are completely different, such as the one shown in Figure 7. In this example, the owner of the box (poss) and a box modifier (mod) were omitted.

This is only his box	**Esta é a caixa**
(b / box	(c / caixa
:poss (h / he)	:domain (e / esta))
:domain (t / this)	
:mod (o / only))	

Figure 7: An example of translation difference

Other differences are language-specific aspects such as the particle “*se*”, a multifunctional word in Portuguese (which, e.g., may represent the conditional “if” or a reflexive pronoun), words that change their part of speech tags and/or are joined in only one word, and other syntactic features. Figures 8 and 9 illustrate some cases. In Figure 8, one may see that the noun “sweetness” becomes the overall concept “sweet-05”, whereas in Portuguese the overall concept is the verb “*rir-01*” (“to laugh”, in English). Moreover, in Portuguese annotation, it is added the *:manner* relation and the “*docemente*” concept (corresponding to “sweetness”).
In Figure 9, the annotation in Portuguese was very different from the English version. Several concepts and relations were left out in Portuguese annotation, for example, the concepts “contrast-01”, “say-01”, “oh” and the relations “:mod” and “:ARG0-of” were omitted in Portuguese annotation. Moreover, we added the “:poss” relation in Portuguese annotation.
Aiming to organize the number of some of these occurrences/phenomena, we computed and summarized them in Table 12. It is important to notice that the hidden subject phenomenon does not change the original annotation, as we make them explicit. An indeterminate subject, on the other hand, is another type of subject (that may include

And there is sweetness in the laughter of all the stars.

```
(a / and
      :op2 (s / sweet-05
            :ARG1 (l / laugh-01
                  :ARG0 (s1 / star
                        :mod (a2 / all)))))
```

E todas as estrelas riem docemente

```
(e / e
      :op2 (r / rir-01
            :ARG0 (e1 / estrelas
                  :mod (t / todas))
            :manner (d / docemente)))
```

Figure 8: Syntactic structuring variation

But the little prince could not restrain his admiration : " Oh !

```
(c / contrast-01
    :ARG2 (p2 / possible-01 :polarity -
        :ARG1 (r / restrain-01
            :ARG0 (p / prince
                :mod (l / little)
                :ARG0-of (s / say-01
                    :ARG1 (o / oh :mode "expressive")))
            :ARG1 (a / admire-01
                :ARG0 p))))
```

O principezinho, então, não pôde conter o seu espanto

```
(p / poder-201 :polarity -
      :ARG0 (p1 / principezinho)
      :ARG1 (c / conter-02
            :ARG0 p1
            :ARG1 (e / espanto
                  :poss p1)))
```

Figure 9: Syntactic structuring variation

the particle "*se*") that may result in changes in the original annotation. The same happens for some different translations, mainly when they incorporate language specific expressions and constructions.

Phenomenon	#	%
Different translation	494	32.35
Syntactic variation	341	22.33
Hidden subject	285	18.66
Missing verb or sense	191	12.50
Change of predicate	100	6.54
Indeterminate subject	68	4.45
Complex predicate	3	0.19

Table 12: Annotation features in Portuguese

In addition to these phenomena, we calculated the incidence of syntactic variations, changes in predication, and missing verbs or senses. Syntactic variations include part of speech changes, as "little prince" (noun-adjective) to "*principezinho*" (noun), "grown-ups" (noun) to "*pessoas grandes*" (noun-adjective), and "boa constrictor" (noun-noun) to "*jibóia*" (noun), among others. Change of predicate occurs when the predicate in Portuguese is different from English. Thus, this change may produce different arguments.

We also computed the number of included arguments (25) and excluded arguments (103) in relation to English. It is also important to add that *VerboBrasil* is still a small dataset compared to PropBank, and, therefore, did not contain all verbs and senses. In cases where the verbs were not in the dataset, we assigned the sense "01" to the verbs, and marked them in the corpus in order to subsidize future improvements in the *VerboBrasil* repository. These cases occurred 191 times.

A final interesting issue is that importing the AMR structures from the English annotation is helpful, but still demands some effort due to the language specificities. As an illustration, each sentence in Portuguese has 8.31 words in average, and we took about 6 minutes to annotate each one, which is less than the English original annotation from scratch, but is still expensive.

5. Final remarks

The annotated corpus should be made available soon, as the Little Prince book went into public domain. We expect that such annotation may foster research in semantic parsing for Portuguese. Our next steps include to perform wikification of the words, as this also happened for English and looks as a natural step to follow.

More than the annotated corpus availability, our contributions are the proposal of an alignment-based approach for AMR annotation, which we believe that may also be used for other language pairs, and the investigation of annotation issues that may be language specific (in spite of the fact of AMR being a meaning representation).

Acknowledgments

The authors are grateful to FAPESP, CAPES, and *Instituto Federal do Piauí* for supporting this work.

References

Banarescu, L., Bonial, C., Cai, S., Georgescu, M., Griffitt, K., Hermjakob, U., Knight, K., Palmer, M., and Schneider, N. (2013). Abstract meaning representation for sembanking. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*, pages 178–186.

Bos, J. (2016). Expressive power of abstract meaning representations. *Computational Linguistics*, pages 527–535.

Burns, G. A., Hermjakob, U., and Ambite, J. L. (2016). Abstract meaning representations as linked data. In *Proceedings of the 15th International Semantic Web Conference*, pages 12–20.

Cai, S. and Knight, K. (2013). Smatch: an evaluation metric for semantic feature structures. In *Proceedings of the*

51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pages 748–752.

Caseli, H. and Nunes, M. (2003). Sentence alignment of brazilian portuguese and english parallel texts. In *Proceedings of the Argentine Symposium on Artificial Intelligence*, pages 1–11.

Damonte, M., Cohen, S. B., and Satta, G. (2017). An incremental parser for abstract meaning representation. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 1, Long Papers*, pages 536–546.

Duran, M. S. and Aluísio, S. M. (2012). Propbank-br: a brazilian treebank annotated with semantic role labels. In *Proceedings of the 8th international conference on Language Resources and Evaluation*, pages 1862–1867.

Duran, M. S., Martins, J. P., and Aluísio, S. M. (2013). Um repositório de verbos para a anotação de papéis semânticos disponível na web. In *Proceedings of the 9th Brazilian Symposium in Information and Human Language Technology*, pages 168–172.

Flanigan, J., Thomson, S., Carbonell, J. G., Dyer, C., and Smith, N. A. (2014). A discriminative graph-based parser for the abstract meaning representation. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, pages 1426–1436.

Goodman, J., Vlachos, A., and Naradowsky, J. (2016). Noise reduction and targeted exploration in imitation learning for abstract meaning representation parsing. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 1–11.

Jurafsky, D. and Martin, J. (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition*. Prentice Hall.

Kingsbury, P. and Palmer, M. (2002). From treebank to propbank. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, pages 1989–1993.

Lehmann, F. (1992). *Semantic networks in artificial intelligence*. Elsevier Science Inc.

Liu, F., Flanigan, J., Thomson, S., Sadeh, N., and Smith, N. A. (2015). Toward abstractive summarization using semantic representations. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1077–1086.

Matthiessen, C. and Bateman, J. A. (1991). *Text generation and systemic-functional linguistics: experiences from English and Japanese*. Pinter Publishers.

Mitra, A. and Baral, C. (2016). Addressing a question answering challenge by combining statistical methods with inductive rule learning and reasoning. In *Proceedings of the 30th Conference on Artificial Intelligence*, pages 2779–2785.

Oepen, S., Kuhlmann, M., Miyao, Y., Zeman, D., Cinkova, S., Flickinger, D., Hajic, J., Ivanova, A., and Uresova, Z. (2016). Towards comparability of linguistic graph banks for semantic parsing. In *Proceedings of the 10th International Conference on Language Resources and Evaluation*, pages 3991–3995.

Ovchinnikova, E. (2012). *Integration of World Knowledge for Natural Language Understanding*. Atlantis Thinking Machines. Atlantis Press.

Palmer, M., Gildea, D., and Kingsbury, P. (2005). The proposition bank: An annotated corpus of semantic roles. *Computational linguistics*, 31(1):71–106.

Pan, X., Cassidy, T., Hermjakob, U., Ji, H., and Knight, K. (2015). Unsupervised entity linking with abstract meaning representation. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 1130–1139.

Peng, X., Song, L., and Gildea, D. (2015). A synchronous hyperedge replacement grammar based approach for amr parsing. In *Proceedings of the 9th Conference on Computational Language Learning*, pages 32–41.

Pourdamghani, N., Knight, K., and Hermjakob, U. (2016). Generating english from abstract meaning representations. In *Proceedings of the 9th International Conference on Natural Language Generation*, pages 21–25.

Song, L., Peng, X., Zhang, Y., Wang, Z., and Gildea, D. (2017). Amr-to-text generation with synchronous node replacement grammar. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 7–13.

Uchida, H., Zhu, M., and Della Senta, T. (1996). Unl: Universal networking language–an electronic language for communication, understanding, and collaboration. *Tokyo: UNU/IAS/UNL Center*.

Wang, C., Xue, N., Pradhan, S., and Pradhan, S. (2015). A transition-based algorithm for amr parsing. In *Proceedings of the 2015 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 366–375.

Zhou, J., Xu, F., Uszkoreit, H., Qu, W., Li, R., and Gu, Y. (2016). Amr parsing with an incremental joint model. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 680–689.