Creation of Latvian FrameNet based on Universal Dependencies

Normunds Gruzitis, Gunta Nespore-Berzkalne, Baiba Saulite
University of Latvia, Institute of Mathematics and Computer Science, Raina blvd. 29, Riga, Latvia
{normunds.gruzitis,gunta.nespore,baiba.valkovska}@lumii.lv

Abstract

This paper presents a work in progress, creating a FrameNet-annotated text corpus for Latvian. This is a part of a larger project which aims at the creation of a multilayered corpus, anchored in cross-lingual state-of-the-art syntactic and semantic representations: Universal Dependencies (UD), FrameNet and PropBank, as well as Abstract Meaning Representation (AMR). For annotating the FrameNet layer, we use the latest frame inventory of Berkeley FrameNet, while the annotation itself is done on top of the underlying UD layer. Thus, the annotation of frames and frame elements is guided by the dependency structure of a sentence, instead of the phrase structure. We strictly follow a corpus-driven approach, meaning that lexical units in Latvian FrameNet are created only based on the annotated corpus examples. Since we are aiming at a medium-sized still general-purpose corpus for a less-resourced language, an important aspect that we take into account is the variety and balance of the corpus in terms of genres, domains and lexical units.

Keywords: FrameNet, Universal Dependencies, Latvian

1. Introduction

Natural language understanding (NLU) systems rely, explicitly or implicitly, on syntactic and semantic parsing of text. State-of-the-art parsers, in turn, typically rely on supervised machine learning which requires substantial language resources – syntactically and semantically annotated text corpora, and extensive linked lexicons.

In the industry-oriented research project “Full Stack of Language Resources for Natural Language Understanding and Generation in Latvian” (Gruzitis et al., 2018), we are creating a balanced text corpus with multilayered annotations, adopting widely acknowledged and cross-lingually applicable representations: Universal Dependencies (UD) (Nivre et al., 2016), FrameNet (Fillmore et al., 2003), PropBank (Palmer et al., 2005) and Abstract Meaning Representation (AMR) (Banarescu et al., 2013).

The UD representation is automatically derived from a more elaborated manually annotated hybrid dependency-constituency representation (Pretkalnina et al., 2016). This also ensures that paragraphs, sentences and tokens are correctly and uniformly split and represented in the standard CoNLL-U data format (see Table 1) before the FrameNet annotation begins. All the annotation layers are afterwards merged, based on the document, paragraph, sentence and token identifiers. The FrameNet annotations are manually added, guided by the underlying UD annotations (see Figure 1). Consequently, frame elements are represented by the root nodes of the respective subtrees instead of text spans; the spans can be easily calculated from the subtrees. The PropBank layer is automatically derived from the FrameNet and UD annotations, provided a manual mapping from lexical units in FrameNet to PropBank frames, and a mapping from FrameNet frame elements to PropBank semantic roles for the given pair of FrameNet and PropBank frames. Draft AMR graphs are to be derived from the UD and PropBank layers, as well as auxiliary layers containing named entity and coreference annotation, with the potential to seamlessly integrate the FrameNet frames and frame elements into the AMR graphs. The semantically richer FrameNet annotations (compared to PropBank) are also helpful in acquiring more accurate draft AMR graphs, even if FrameNet itself stays behind the scenes.

The inspiration to create an integrated multilayer corpus comes from the OntoNotes corpus (Hovy et al., 2006) and the Groningen Meaning Bank (GMB) (Bos et al., 2017). The overall difference from the OntoNotes approach is that we use the UD model at the treebank layer, and we annotate FrameNet frames in addition to the PropBank frames. In fact, FrameNet is the primary frame-semantic representation in our approach. Another difference is that we aim at whole-sentence semantic annotation at the ultimate AMR layer. This in some sense is similar to the goal of GMB, but GMB uses Discourse Representation Theory instead of AMR. For pragmatic reasons, we use the more shallow and more lossy AMR formalism. Our experience developing semantic parsers and multilingual text generators, by combining machine learning and grammar engineering (Gruzitis et al., 2017; Gruzitis and Dannells, 2017), has convinced us that FrameNet and AMR both have a great potential to establish as powerful and complementary semantic interlinguas which can be furthermore strengthened and complemented by other multilingual frameworks, like Grammatical Framework (Ranta, 2011).

In this paper, we focus on the creation of the intermediate FrameNet layer of the full-stack multilayer corpus. Note that the current project addresses only frequently used verbs as frame-evoking lexical units. A spin-off project has been just launched to work on frequent nominalizations, following the same methodology.

It should also be noted that there has been previous work on a domain-specific Latvian FrameNet for a real life media monitoring use case, focusing on 26 modified Berkeley FrameNet (BFN) frames (Barzdins et al., 2014). The current work, however, aims at a balanced general-purpose BFN-compliant framenet that will cover many frequently used frames and lexical units.

Although this paper focuses on Latvian, we believe that our experience and findings can be useful for the creation of dependency treebank based framenets for other less-resourced languages as well.

Table 1: A sample sentence *"On Wednesday evening, the nation's beloved poet Imants Ziedonis passed away at age 79."* represented in the CoNLL-U data format. Field FEATS is left empty because of space restrictions. The literal translations are not part of CoNLL-U.

ID	FORM	LEMMA		UPOSTAG	XPOSTAG	FEATS	HEAD	DEPREL	DEPS
1	Trešdienas	trešdiena	'Wednesday'	NOUN	ncfsg4		2	nmod	2:nmod:gen
2	vakarā	vakars	'evening'	NOUN	ncmsl1		7	obl	7:obl:loc
3	79	79	'79'	NUM	xn		4	nummod	4:nummod
4	gadu	gads	'year'	NOUN	ncmpg1		5	nmod	5:nmod:gen
5	vecumā	vecums	'age'	NOUN	ncmsl1		7	obl	7:obl:loc
6	mūžībā	mūžība	'eternity'	NOUN	ncfsl4		7	obl	7:obl:loc
7	aizgājis	aiziet	'leave'	VERB	vmnpdmsnasnpn		0	root	0:root
8	tautā	tauta	'nation'	NOUN	ncfsl4		9	obl	9:obl:loc
9	mīlētais	mīlēt	'love'	VERB	vmnpdmsnpsypn		10	amod	10:amod
10	dzejnieks	dzejnieks	'poet'	NOUN	ncmsn1		11	nmod	11:nmod
11	Imants	Imants	'Imants'	PROPN	npmsn1		7	nsubj	7:nsubj
12	Ziedonis	Ziedonis	'Ziedonis'	PROPN	npmsn2		11	flat:name	11:flat:name
13	.	.	'.'	PUNCT	zs		7	punct	7:punct

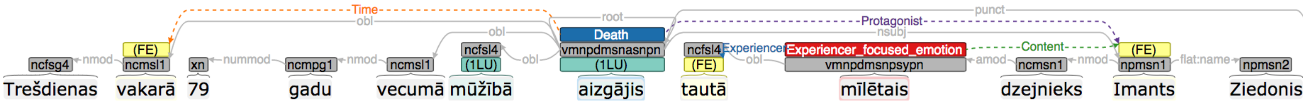

Figure 1: FrameNet annotation in WebAnno on top of a UD tree (Table 1). Only head nodes are selected while annotating frame elements (FE). The FE spans can be acquired automatically by traversing the respective subtrees: [*trešdienas* *vakarā*]$_{\text{Time}}$, [*tautā*]$_{\text{Experiencer}}$, [*tautā mīlētais dzejnieks Imants Ziedonis*]$_{\text{Protagonist}}$. Multi-word lexical units (LU) are indicated by generic LU tags: *mūžībā aizgājis*$_{\text{DEATH}}$ versus *mīlētais*$_{\text{EXPERIENCER\_FOCUSED\_EMOTION}}$.

2. The corpus

In this project, we are aiming at a medium-sized corpus – around 10,000 sentences annotated at all the layers mentioned in Section 1. Therefore it is crucially important to ensure that the multilayer corpus is balanced not only in terms of text genres and writing styles but also in terms of lexical units.

Our fundamental design decision is that the text unit is an isolated paragraph. The corpus therefore consists of manually selected paragraphs from many different texts of various types. Representative paragraphs are selected in different proportions from a balanced 10-million-word text corpus: around 60% come from various news sources, around 20% is fiction, around 10% are legal texts, around 5% is spoken language (parliament transcripts), and the rest is miscellaneous.

As for the lexical units, our goal is to cover at least 1,000 most frequently occurring verbs, calculated from the 10-million-word corpus. Since the most frequent verbs tend to be also the most polysemous, we expect that the number of lexical units (verb senses w.r.t. FrameNet frames) will be considerably larger – at least 1,500 units. At this stage, it is too early to predict any numbers regarding nominal lexical units. Nevertheless, the more frequent a lexical unit is, the more annotated examples it will have. We are aiming at around 10 annotation sets per lexical unit on average.

Paragraphs to be annotated are selected based on verbs they contain, not randomly, and curators are constantly updated on the current balance or imbalance of the corpus w.r.t. genres and verb frequencies. We assume that the corpus will turn out to be balances also w.r.t. nominal lexical units.

Our decision about the data selection is justified also by the lessons learned in other treebanking and framebanking projects. For instance, Bick (2017) concludes that a sentence-randomized propbank not only has a limited usage for coreference resolution and discourse analysis but also provides a limited coverage of lexical units.

At the time of writing, we have acquired more than 5,000 annotation sets (by investing around four man-months). This data set already covers more than 300 BFN frames evoked by nearly 900 lexical units.

The Latvian FrameNet corpus is being gradually released on GitHub under the CC BY-NC-SA 4.0 license.[1]

3. The FrameNet annotation process

Paragraphs for which the manual treebank annotation is finalized and which have been successfully converted to the UD representation are considered for the FrameNet annotation. Unfinished paragraphs are ignored till next iteration, since their sentence split, tokenization, as well as tree structure can still considerably change. Changes in the tree structure is not a major issue, and the FrameNet annotation process actually helps to spot and eliminate many inconsistencies in the underlying trees. The sentence splitting and tokenization, however, is a major requirement to later avoid issues in merging the different annotation layers.

Since we annotate FrameNet frames on top of UD trees, we need an annotation tool which supports both representations. Therefore we have chosen the WebAnno platform (Eckart de Castilho et al., 2016) which also supports a centralized web-based annotation workflow.

[1] https://github.com/LUMII-AILab/FullStack

3.1. The concordance approach

While treebank, named entity and coreference annotations are done paragraph by paragraph and sentence by sentence, we do not find this being a productive workflow for annotating semantic frames, especially in case of the highly abstract FrameNet frames. Instead, we prefer a concordance view, so that the linguist can focus on a target word and its different senses (frames), without constantly switching among different sets of frames. This also improves the annotation consistency.

To provide such annotation environment, we automatically extract all UD-annotated sentences from the finalized paragraphs containing the requested target word, and we store the result in a separate CoNLL-U file. More precisely, we group sentences for FrameNet annotation by applying filters on the LEMMA and POSTAG fields in the CoNLL-U files (see Table 1), as well as the DEPREL field in case of nominalizations (e.g. participles having the *amod* or *nmod* dependency relation).

Figure 2 illustrates a partial concordance with FrameNet annotations. The UD annotations are hidden for the sake of simplicity, and, in fact, they are hidden also in the curation view in WebAnno.[2] The actual annotation, however, is done on top of the UD layer, as illustrated in Figure 1.

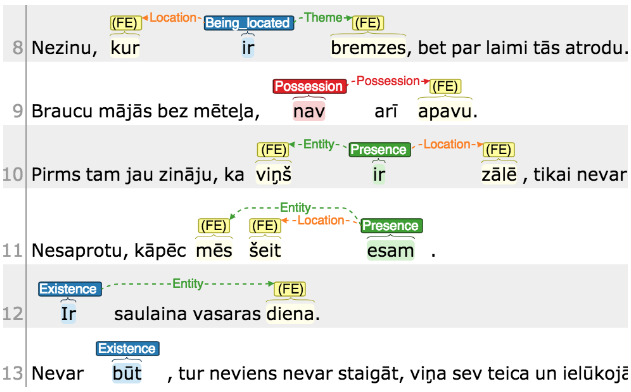

Figure 2: A screenshot of the WebAnno tool: FrameNet-annotated occurrences of the target verb *būt* ('to be located', 'to be present', 'to have', or 'to exist').

When more paragraphs are finalized at the UD layer, they are included in the next concordance queries. In practice, for each target word there will be at least two concordance files extracted and annotated during the project. The first concordance is processed when there are at least three example sentences available for the target word. The second concordance will be processed when the planned 10,000 sentences will be done at the UD layer. The second concordance will contain only the new examples which are not included in the first concordance (according to the sentence identifiers). The annotated concordances from the first round will serve as guidelines when annotating the second round, thus, further improving consistency.

A consequence of such approach is that no full-text annotation is intentionally done, although many sentences might become fully or almost fully annotated after merging annotations of the same sentence from different concordances (see Table 2).

3.2. The UD-based annotation

The acquired UD-annotated concordances (full sentences) are imported in the WebAnno platform which we have specifically configured for the FrameNet annotation. To facilitate the annotation process, we have generated two kinds of WebAnno constraint sets. First, a set of frame to core frame element mappings (from BFN 1.7 data), so that a menu of core frame elements is generated when the annotator selects a frame for the particular occurrence of the target word. Second, a set of LEMMA/POSTAG to frame mappings, so that the most probable frames (senses) for the particular occurrence of the target word are highlighted at the top of the frame selection menu.

The UD-based approach has a significant consequence: frame elements are not annotated as spans of text – annotators select only the head word (node) when annotating a frame element. The whole span can be easily calculated automatically by traversing the respective UD subtree. These calculations are not included as part of the data set.

Such approach not only makes the annotation process more simple and the annotations more consistent, but it also facilitates the training of an automatic semantic role labeler, since it is easier to identify the syntactic head of a frame element than a span of a string. Still, most FrameNet corpora are annotated in terms of spans, relying on syntactic parsing as a post-processing step.

When the FrameNet annotation is done, the finalized concordances are exported from WebAnno, and are converted from the TSV3 format used by WebAnno to a more common CoNLL 2009-like format which combines the UD and FrameNet annotations (see Table 2). During conversion, the UD data fields in the CoNLL-2009 output are updated from the latest version of the UD treebank, and the isolated sentences are eventually reorganized back into paragraphs.

3.3. Important notes on frame elements

Yet another important decision regarding frame elements is to annotate only the core elements according to BFN. We have made this decision because of the limited time frame and the wider scope of the current project. However, we do annotate two non-core elements systematically: *Time* and *Place* (as illustrated in Figure 1). Our industrial partner, the national news agency LETA, is interested in the automation of media monitoring processes. In their information extraction use case, these two non-core frame elements are important, and they will be informative in other use cases as well. Other non-core elements are annotated occasionally, if they are rather specific to the frame (e.g. non-core indirect objects and specific adverbial modifiers).

Regarding null instantiations (NI), we do not annotate missing frame elements in the sentence. This is out of the scope of the current project, but the annotation of NI should to be considered in a follow-up research: (i) since the FrameNet annotation is relaying on UD, it is an open question how to handle NI – where to attach these annotations; (ii) since Latvian is a highly inflected language, the grammatical sub-

[2] Each concordance is annotated by one linguist and curated by another linguist, which is supported in WebAnno.

Table 2: A data format used to serialize the FrameNet layer of the corpus: a version of CoNLL-2009 based on CoNLL-U (see Table 1). Several CoNLL-U fields are excluded from this table because of space restrictions.

ID	FORM	LEMMA	UPOSTAG	XPOSTAG	DEPS	**FILLPRED**	**PRED**	**APRED$_1$**	**APRED$_2$**
1	Trešdienas	trešdiena	NOUN	ncfsg4	2:nmod:gen	_	_	_	_
2	vakarā	vakars	NOUN	ncmsl1	7:obl:loc	_	_	Time	_
3	79	79	NUM	xn	4:nummod	_	_	_	_
4	gadu	gads	NOUN	ncmpg1	5:nmod:gen	_	_	_	_
5	vecumā	vecums	NOUN	ncmsl1	7:obl:loc	_	_	_	_
6	mūžībā	mūžība	NOUN	ncfsl4	7:obl:loc	_	_	_	_
7	aizgājis	aiziet	VERB	vmnpdmsnasnpn	0:root	Y	Death	_	_
8	tautā	tauta	NOUN	ncfsl4	9:obl:loc	_	_	_	Experiencer
9	mīlētais	mīlēt	VERB	vmnpdmsnpsypn	10:amod	Y	Experiencer_focused_emotion	_	_
10	dzejnieks	dzejnieks	NOUN	ncmsn1	11:nmod	_	_	_	_
11	Imants	Imants	PROPN	npmsn1	7:nsubj	_	_	Protagonist	_
12	Ziedonis	Ziedonis	PROPN	npmsn2	11:flat:name	_	_	_	_
13	.	.	PUNCT	zs	7:punct	_	_	_	_

ject and object can be omitted in a sentence, to some extent, compensating it with the respective form of the verb; (iii) in general, it would require Latvian-specific guidelines, but the theoretical foundations are not mature yet for Latvian; it would require more elaborate linguistic research, based on the basic annotated data acquired in the current project; (iv) although NI is highly relevant for lexicographic research, it is not a priority for many practical use cases that require semantic parsing.

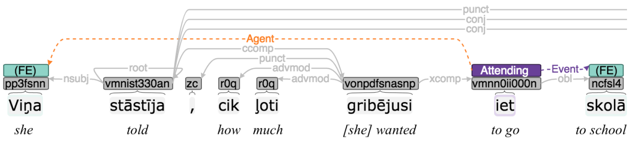

Figure 3: Non-projective annotation of a frame element (FE): the frame *Attending* is evoked in a subclause while its FE *Agent* is mentioned in the main clause.

It should be noted, however, that we do annotate frame elements that non-projective w.r.t. the underlying UD tree structure, i.e., that syntactically are not arguments of the target verb. Figure 3 provides an example.

3.4. Multi-word lexical units

Regarding lexical units, although we focus on verbs, they some times must be considered as multi-word units or constructions. To deal with this issue, we have introduced an auxiliary annotation layer for multi-word lexical units (as illustrated in Figure 1). The head word is still a verb that evokes a frame, but the other key constituents are indicated as well. Again, note that these constituents are roots of the respective subtrees (in general) – we do not annotate the whole spans.

This auxiliary layer is not an ultimate solution to deal with constructions, but for now it allows us to register such cases and to retrieve them later for more elaborated analysis. Usually these are partially grammaticalized constructions or even idioms that, as a whole, evoke the respective frames. If we would consider these verbs in isolation, they would rather evoke different frames, e.g.:

iet bojā ‘to die’ (*iet* – ‘to go’);

aiziet mūžībā ‘to pass away’ (*aiziet* – ‘to leave’);

ņemt vērā ‘to consider’ (*ņemt* – ‘to take’);

nākt klajā ‘to be published’ (*nākt* – ‘to come’);

nākt par labu ‘to be beneficial’ (*nākt* – ‘to come’);

likt lietā ‘to use’ (*likt* – ‘to put’).

3.5. Cross-lingual issues

In order to ensure compliance with the Berkeley FrameNet and, thus, to maximize the cross-lingual applicability of Latvian FrameNet, we are strictly sticking to the BFN frame inventory. We avoid defining any Latvian-specific frames. Therefore it is sometimes difficult to select an appropriate BFN frame for a particular sense of a Latvian verb. It usually happens when:

1. The sense of a Latvian verb is more specific compared to the closest English verb sense or compared to the definition of the closest BFN frame. For instance, for the verb *pārdomāt* ‘to change one’s mind’ or ‘to rethink’, we do not have a solution yet, since BFN frames related to thinking (*Opinion*, *Cogitation*) do not fit this verb sense, and neither does the general *Cause_change* frame. Similarly, we have not found a good mapping for *maldīties* ‘to be wrong’ and *saņemties* ‘to pull oneself together’.

2. The sense of a Latvian verb is more general compared to the closest English verb sense: the sense of an English verb is expressed in Latvian by a phrase (typically, by a verb and a direct object). Examples: *lasīt lekciju* ‘to lecture’ (‘to give a lecture’), *krist ģībonī* ‘to faint’ (‘to fall into unconsciousness’), *zaudēt samaņu* ‘to faint’ (‘to lose consciousness’).

3. The semantic elements are different between the Latvian and English verb senses. For instance, *braukt* ‘to move using a vehicle’: the sense of the Latvian verb does not specify whether the person is a driver or a passenger (e.g. *es braucu uz darbu* ‘I go to work (by a transport)’ – it is unclear what is the role of the person, and which frame is evoked – *Ride_vehicle* or *Operate_vehicle*. In this particular case, we use the frame *Use_vehicle* which is a non-lexical frame in English.

There are some options how to deal with these issues: (i) by treating more verb phrases in Latvian as if they were multi-word lexical units, even if lexicographers would argue about that (the second point in the above listing); (ii) by using a more general BFN frame if possible, i.e., if the direct object of the target verb can be annotated as a core frame element (e.g., it would work for 'to lose consciousness' but not for 'to give a lecture'); (iii) some frames are just missing in BFN, and a global solution would be needed on how to propose and confirm new frames in the BFN frame hierarchy; most likely in the scope of the Multilingual FrameNet initiative (Gilardi and Baker, 2018).

4. Conclusion

Creating the Latvian FrameNet, we strictly follow a corpus-driven approach: no lexical units are introduced without annotated examples, i.e., we create no lexical units based on lexicographic intuition or a common-sense dictionary; only based on corpus evidence. An initial experiment on bootstrapping lexical units without corpus evidence did not prove to be productive: many of those hypothesis are not confirmed by our corpus (at least for now), and vice versa – many lexical units were missing.

The consecutive treebank and framebank annotation workflow has turned out very productive and mutually beneficial. The dependency tree facilitates the annotation of semantic frames and roles, while the frame semantic analysis of the verb valency often unveils various inconsistencies and bugs in the dependency or morphological annotation. These issues are immediately fixed in the treebank, and are later automatically synchronized with the FrameNet layer.

Because of the UD-based approach, we cannot use the specialized annotation tools developed for Berkeley FrameNet, or FrameNet Brasil, for instance. However, conversion to the BFN data format (from a CoNLL-like format) is possible (by using UD dependency relations instead of phrase types, etc.), so that the BFN-compliant web tools could be used at least for viewing and browsing Latvian FrameNet.

5. Acknowledgements

This work has received financial support from the European Regional Development Fund under the grant agreements No. 1.1.1.1/16/A/219 and No. 1.1.1.2/VIAA/1/16/188.

6. Bibliographical References

Banarescu, L., Bonial, C., Cai, S., Georgescu, M., Griffitt, K., Hermjakob, U., Knight, K., Koehn, P., 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, Sofia, Bulgaria.

Barzdins, G., Gosko, D., Rituma, L., and Paikens, P. (2014). Using C5.0 and exhaustive search for boosting frame-semantic parsing accuracy. In *Proceedings of the 9th International Conference on Language Resources and Evaluation (LREC)*, pages 4476–4482, Reykjavik, Iceland.

Bick, E. (2017). From Treebank to Propbank: A Semantic-Role and VerbNet Corpus for Danish. In *Proceedings of the 21st Nordic Conference on Computational Linguistics*, pages 202–210, Gothenburg, Sweden.

Bos, J., Basile, V., Evang, K., Venhuizen, N., and Bjerva, J. (2017). The Groningen Meaning Bank. In Nancy Ide et al., editors, *Handbook of Linguistic Annotation*, volume 2, pages 463–496. Springer.

Eckart de Castilho, R., Mújdricza-Maydt, E., Yimam, S. M., Hartmann, S., Gurevych, I., Frank, A., and Biemann, C. (2016). A Web-based Tool for the Integrated Annotation of Semantic and Syntactic Structures. In *Proceedings of the Workshop on Language Technology Resources and Tools for Digital Humanities*, pages 76–84, Osaka, Japan.

Fillmore, C. J., Johnson, C. R., and Petruck, M. R. (2003). Background to FrameNet. *International Journal of Lexicography*, 16(3):235–250.

Gilardi, L. and Baker, C. (2018). Learning to Align across Languages: Toward Multilingual FrameNet. In *International FrameNet Workshop 2018: Multilingual FrameNets and Constructicons*, Miyazaki, Japan.

Gruzitis, N. and Dannells, D. (2017). A multilingual FrameNet-based grammar and lexicon for Controlled Natural Language. *Language Resources and Evaluation*, 51(1):37–66.

Gruzitis, N., Gosko, D., and Barzdins, G. (2017). RIGOTRIO at SemEval-2017 Task 9: Combining machine learning and grammar engineering for AMR parsing and generation. In *Proceedings of the 11th International Workshop on Semantic Evaluation*, pages 924–928, Vancouver, Canada.

Gruzitis, N., Pretkalnina, L., Saulite, B., Rituma, L., Nespore-Berzkalne, G., Znotins, A., and Paikens, P. (2018). Creation of a Balanced State-of-the-Art Multilayer Corpus for NLU. In *Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC)*, Miyazaki, Japan.

Hovy, E., Marcus, M., Palmer, M., Ramshaw, L., and Weischedel, R. (2006). OntoNotes: The 90% Solution. In *Proceedings of the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers*, pages 57–60, New York City, USA.

Nivre, J., de Marneffe, M.-C., Ginter, F., Goldberg, Y., Hajič, J., Manning, C. D., McDonald, R., Petrov, S., Pyysalo, S., Silveira, N., Tsarfaty, R., and Zeman, D. (2016). Universal Dependencies v1: A Multilingual Treebank Collection. In *Proceedings of the 10th International Conference on Language Resources and Evaluation*, pages 1659–1666.

Palmer, M., Gildea, D., and Kingsbury, P. (2005). The Proposition Bank: An Annotated Corpus of Semantic Roles. *Computational Linguistics*, 31(1):71–106.

Pretkalnina, L., Rituma, L., and Saulite, B. (2016). Universal Dependency treebank for Latvian: A pilot. In *Human Language Technologies – The Baltic Perspective*, volume 289 of *Frontiers in Artificial Intelligence and Applications*, pages 136–143. IOS Press.

Ranta, A. (2011). *Grammatical Framework: Programming with Multilingual Grammars*. CSLI Publications, Stanford.