

Welcome to Release 1.7 of the FrameNet data

We are happy to announce Release 1.7 of the FrameNet lexical database. A more than one year has passed since Release 1.6, and we are trying to make these data releases more frequent. Although the data on the FrameNet public website is updated on a nightly basis, we believe that a more formal data release from time to time is needed, both to provide a standard for benchmarking software that uses the FrameNet data, and as an occasion to tell users of the FrameNet data about recent changes.

We are pleased at the increasing use of FrameNet as a resource for semantic analysis both by academic and commercial users. We are again licensing Release 1.7 of the FrameNet data under a Creative Commons Attribution-Only license. This license allows use for any purpose, including commercial use, and requires public acknowledgment of all such uses. In general, academic users have included clear acknowledgments in published papers, but we are aware of a number of commercial users who have failed to do so. If you produce a product or provide a service based on the use of the FrameNet data, **you are required by the license agreement** to provide your users with a written reference to FrameNet, including the URL of the FrameNet website (<http://framenet.icsi.berkeley.edu>) and a reference to a current FrameNet publication. (The most complete, reasonably current published summary of work on FrameNet is Fillmore and Baker (2010) “A Frame Semantic Approach to Linguistic Analysis” in Heine and Narrog (eds.) *Oxford Handbook of Linguistic Analysis*.)

1 Growth of the FrameNet database

This data release does not cover as long a time span or include as large a jump in the numbers of frames and lexical units (LUs) as did Release 1.6, in part because of lower funding levels. However, a number of conceptually and linguistically important frames have been added, such as three frames based on Force dynamics (cf. Talmy (2000)). Many changes to the frame relations have also been made, especially in the upper levels of the hierarchy, relating to events and causation; we believe these will make the database more useful for reasoning about events. These changes are discussed below.

Table 1 gives some statistics on each release of the data.¹ In all, we have added 19 frames, deleted 3 and renamed 16. At the Frame Element (role) level, we have added 181 Frame Elements (FEs), deleted 3, and renamed 21. The deletions and many of the renamings have to do with the revised treatment of Force Dynamic frames, which is discussed below. The numbers reported in this table are very close to correct, but were compiled according to several different principles and may not add exactly to totals shown on the FrameNet public website “Current project status” page.

The file “DifferencesR1.6-R1.7.xml” provides detailed information on changes to the frames, frame elements, and lexical units in this release, including which frame element changes are due to adding or deleting frames, and which are additions to or deletions from existing frames. Note that the IDs of frames, frame elements, and lexical units do not change when they are renamed, so wherever possible, any comparisons or mappings across versions of the FrameNet data should link via ID numbers rather than names.

At the document level, we added three documents under “Miscellaneous” and renamed the corpus “WikiNews” to “WikiTexts”, because we also added seven files which represent

¹A version called Release 1.4 was begun, but never officially distributed.

experiments with lexicographic annotation (i.e. one lemma per clause) on text extracted from Wikipedia. Finally, there were four documents in Release 1.6 containing a small amount of full text annotation which were almost duplicates: ANC_journal_christine, ANC_journal_patrick, ANC_journal_ryan, and ANC_journal.pbio.0020001; we kept the first of these and removed the other three from Release 1.7.

	R1.2	R1.3	R1.5	R1.6	R1.7
Frames	609	795	1,019	1,205	1,221
(non-lexical frames)	58	74	111	134	135
FEs in lexical frames	4,909	7,124	8,884	10,333	10,503
FE/lexical frame	8.91	9.88	9.78	9.65	9.7
Proportion non-lexical	9.5%	9.3%	10.9%	11.1%	11.1%
Frame relations	550	1,152	1,507	1,805	2,070
FE relations	2,770	6,311	8,252	11,791	12,393
Lexical Units	8,869	10,195	11,829	13,312	13,572
LUs/lexical frame	16.1	14.14	13.03	12.43	12.5
LUs w/ lexicog. anno	6642	6,815	7,711	8,313	8,390
Pct. LUs w/ lexicog. anno	74.9%	66.8%	65.2%	62.5%	62%
AnnoSets in lexicog anno	13,3846	139,439	149,931	157,739	174,017
Lexicog. AnnoSets/annotated LU	20.2	20.5	19.5	19.0	20.73
AnnoSets in full text anno	0	11,671	23,087	28,160	28,208
Total AnnoSets	133,846	151,110	173,018	201,695	202,225
Full Text anno/Total anno	0%	7.7%	13%	22%	14%

Table 1: Comparison of Releases 1.2, 1.3, 1.5, 1.6 and 1.7

For example, we created a new frame `Improvement` or `decline` which includes new Frame Elements such as the `ENTITY` that undergoes the change, and the `DEGREE` of the improvement or decline. We also added 455 new Frame Elements to existing frames, such as the FE `DEGREE` in the existing frame `Notability`, and `TIME` in the frame `Being` contained. These are marked in the file “DifferencesR1.6-R1.7.xml” with the attribute `added_to_frame="y"`.

We have also updated the supplementary XML files containing all the mappings from lexemes and lemmas to word forms in the database. See the Release 1.6 notes for more information on these files. We have also defined a new part-of-speech, “idiosyncratic”, abbreviated “`IDIO`”, for lexemes that appear not to fall into the normal categories; so far it has been used only for the lexeme *ago*.

2 Some major changes in Release 1.7

2.1 Revision of “The Book”

“The Book”, *FrameNet II: Extended Theory and Practice*, which documents the principles and practice of FrameNet annotation, has been extensively revised. Since it has been written over a long time by numerous authors, some sections were out of date, or reflected earlier stages in the development of FrameNet. The current authors have done their best to make sure that it adequately represents the current situation.

2.2 Changes to the upper-level event hierarchy

A number of changes have been made near the top of the hierarchy which describe events related to change. The frames involved include Becoming, Undergo change, Undergo transformation, Transition to a situation, Transition to a quality, and Transition to a state. We have now cleanly separated out three different ways of thinking about change:

1. As a holistic change of the situation, for which we use the FE names INITIAL_SITUATION and FINAL_SITUATION
2. As a change in some quality of an entity, for which we use INITIAL_QUALITY and FINAL_QUALITY
3. As a change in the category of an entity, with FE names like INITIAL_CATEGORY and FINAL_CATEGORY.

The difference among these is exemplified in (1), (2), and (3), respectively.

- (1) They **ended up** [taking the bus FINAL_SITUATION].
- (2) Londoners just **grew** [angrier FINAL_QUALITY].
- (3) The church **became** [a nightclub FINAL_CATEGORY].

The lexical units in this domain, such as *become*, *transform*, *go*, *come*, etc. exhibit a number of different patterns of allowing or disallowing realization of the above FEs. All of the lexical units of this domain allow expression of the final situation, quality, or category, but some focus on the difference between the before and after stages, and allow expression of both initial and final state. In keeping with FrameNet’s policy of frame separation, each alternation pattern has been assigned to a different frame. The patterns are summarized in Table 2.

Frame name	Situation	Quality	Category	Initial OK?	Example LU
Transition to a state	+	+	+	-	<i>end up.v</i>
Becoming	-	+	+	-	<i>become.v</i>
Transition to a quality	-	+	-	-	<i>go.v</i>
Transition to a situation	+	-	-	-	<i>come.v</i>
Undergoing change	+	+	+	+	<i>change.v</i>
Undergo transformation	-	-	+	+	<i>transform.v</i>

Table 2: Frames related to change

Of the listed frames, most have a normal complement of lexical units, but two have only one lexical unit each: Transition to a quality, with *go.v* in cases like “**Go** crazy”, and Transition to a situation with *come.v* in cases like “She **came** to loathe him”. It is not surprising that two of the most basic words in English should have special behavior, and it is appropriate for them to have separate, unique frames.

2.3 Semantic Type Hierarchy for Spatial relations

Having added many frames and semantic types for static spatial relations in R1.6, we have improved this domain both by adding annotation in these frames and by creating inheritance relations among semantic types introduced in R1.6, forming a separate small hierarchy distinct from our frame hierarchy. This semantic type hierarchy represents useful generalizations about the semantics of the frames, such as the fact that NORTH, SOUTH, EAST and WEST are CARDINAL DIRECTIONS, and CARDINAL DIRECTIONS are a subtype of ABSOLUTE DIRECTIONS, along with UP and DOWN. Both of these categories have special syntax and semantics in many languages, such as the well-documented extensive use of Cardinal directions in some Australian languages and the full set of compass directions in English (e.g. the special compounds and constructions such as *north northwest* and *north by northwest*. (See The Book, Section 6.2 for a fuller discussion of FrameNet semantic types.)

2.4 Force Dynamic frames

The three new frames in this group, Level of force exertion, Level of force resistance, and Dynamism, are based on the theory of Force Dynamics (Talmy, 2000), relating to situations that can be understood in terms of the application of force and resistance to that force. They represent different perspectives on the situation and individual LUs in these frames such as *force.n*, *powerful.a*, *weak.a*, *frail.a* *hardy.a*, *sluggish.a*, and *stamina.n* describe different levels of force and resistance from different points of view. For further discussion, see the definition of the frame Level of force exertion in the FrameNet data.

2.5 New Frame-to-frame relation: Metaphor

Metaphor is an important, fundamental feature of everyday language, but FrameNet has long had a somewhat informal representation of it, which is discussed in detail in The Book, Section 6.4 “Annotation of Metaphors in FrameNet”.

The FN team (and Prof. Charles J. Fillmore, the founder of the project) were reluctant to add a Metaphor relation to the database, largely because in most cases, not all the LUs of the source frame participate in the metaphorical mapping. For example, many of the LUs for physical attacking (in the Attack frame) can be used for verbal attacks (in the Judgement communication fame) such as *attack*, *assault*, *snipe at*, but others cannot (*sap*, *small arms fire*, *storm*, etc.). A Metaphor relation between these two frames would give no clue as to which LUs are in which group. The only partially predictive power of such a Metaphor relation is in contrast to relations like Inheritance, where all of the LUs of the child frame are asserted to be instances of the parent frame.

Nevertheless, partially as a result of working with the members of the MetaNet project at ICSI from 2012 to 2015, we decided to make a new frame-to-frame relation called Metaphor in the FrameNet database. We have created only a few instances in the database as of Release 1.7, but we plan to create many more as our resources permit. We believe that this will be an important step toward creating an adequate representation of causation, which is often described in the vocabulary of Force Dynamics even for abstract domains (e.g., *put pressure on her to act*, *resist the forces of change*). Note, however, that this new relation is not currently shown in the frame definitions in the HTML reports, because the XSLT has not been changed to show them, but the new relation is in XML for Release 1.7.

2.6 Modals, Negation, Conditionals, etc.

Expanding on the work on negation and conditionals in Release 1.6, we have connected negation, conditionals, causation and a number of modals into a single connected graph of frames. The main impetus for our structure is the concept of *alternative spaces* in Mental Space Theory, i.e. the basic scene where there are two mutually exclusive ideas about the current or a potential state of the world. (See the high-level frame `Alternativity`.)

- This approach models **negation** in an obvious way, since a sentence like *We didn't eat corn dogs* presents a factual situation as contrasting with a mutually exclusive non-factual situation, *we ate corn dogs*. (See the `Negation` frame.)
- **Conditionals** contrast a situation under a particular condition, e.g. *if we win, ...* against the background of a mutually exclusive situation *if we don't win, ...*; see the frame `Conditional scenario`.
- **Causation** is modeled as profiling a factual situation of the `EFFECT`, given a `CAUSE`. Contrasting with the non-occurring, or `PREVENTED SITUATION` that would have occurred without the `CAUSE`. (See the definition of the `Causation scenario`. Section A.7 in the appendix of The Book, “FEs related to the causal chain” also details recent changes to widely used extrathematic FEs related to causation.)
- **Deontic modals** such as *should* and *must* also present a contrast, between a possibility that would be desirable and a possibility that is either undesirable or unacceptable.

Work on this area was advanced in response to the needs of a commercial user of FrameNet; we hope that other users who wish to use the FN resources for modeling situations in text will find the more complete coverage of these domains useful.

2.7 Support and Control relations

We have clarified and recategorized the labels for control and support relations in a way that improves the distinction between these relations and true multiword expressions (in FrameNet, multi-lexeme lemmas). See Sec. 3.2. of The Book for a full discussion.

3 Planned Improvements

3.1 Frame relations

We will review instances of `Inheritance` and `Using` relations which we created early in the course of the FrameNet Project and determine if some of them are better described as `Perspective on relations`. (See The Book, Section 6.1 for an up-to-date description of all frame relations.)

3.2 Verb particle POS marking

In a number of multiword LUs containing what are traditionally called “verb+particle” combinations, the particle has been erroneously assigned other parts of speech (`Adv` or `Prep`) when they should have had the part of speech `AVP` for “adverbial particle”. We plan to fix these errors in the very near term.

ID	Parent name	ID	Child Name
254	Axis_specified	118	Landmark_based
118	Landmark_based	119	Top
118	Landmark_based	121	Landmark_front
118	Landmark_based	122	Landmark_back
118	Landmark_based	123	Landmark_left
118	Landmark_based	124	Landmark_right
133	Motion_based	134	Ahead
133	Motion_based	135	After
136	Viewpoint_based	137	Viewpoint_left
136	Viewpoint_based	138	Viewpoint_right
116	On_top	130	Up

Table 3: Omitted semantic type inheritance relations

3.3 Multiple inheritance in Semantic type hierarchy

We had planned to add a few more inheritance links between semantic types, which would cause certain types to have more than one parent. However, we found that multiple inheritance of semantic types causes problems for our database server software. Therefore, the inheritance relations shown in Table 3 have been left out of the database for the time being. We hope to revise the software to allow us to insert these missing relations.

3.4 XML format and visualization with XSLT

There have been no changes to the XML format since release 1.6, so all software that works with R1.6 should work with R1.7. However, the new Metaphor relation is still not displayed by the current XSLT scripts; they will need to be revised to show it.

Not all web browsers fully support XSLT scripts, and it is not possible to test all combinations of versions of operating systems and browsers. However, based on machines and software available to the FrameNet team, we can report the following:

- Firefox 45.x or later works correctly for browsing both the FrameNet website and the downloaded Release 1.7 data files on Red Hat Enterprise Linux 6, Mac OS 10.11 (El Capitan), and Windows 7.
- On Mac OS, Google Chrome 54.x and Safari 10.x work correctly on the FrameNet website; Safari also works with the downloaded data out of the box, but for Chrome to work correctly with the downloaded data, it has to be run with the parameter “--allow-file-access-from-files”, which is considered a security risk in some circumstances. (As long as it is only being used on the downloaded FrameNet data, and not connecting to the web, there should be no security risk.)
- All three browsers work for the FrameNet website on iOS 10.1 on the iPhone; we hope to modify the website in the near future to improve the user interface on small screens.
- On Windows 7, Internet Explorer 11 works correctly for the FrameNet website, and will work for the downloaded data if the user gives permission by answering “yes” to the question “allow to run scripts?”.

3.5 Multilingual FrameNet

Finally, as many users of FrameNet will be aware, researchers in a number of countries are building FrameNet-like lexical databases for languages other than English. We have included links to some of these projects on the FrameNet website. We are pleased to announce that we have received an NSF grant (#1629989) entitled “Multilingual FrameNet: A Resource Enabling Cross-Lingual Research for the Natural Language Processing Community”. This grant will support the creation of a database of alignments among FrameNet-related projects in at least seven or eight of these languages. This will aid research into the questions of how similar semantic frames are across languages, to what extent the use of FrameNet frames can improve translation between these languages, how similar valence patterns are across languages, etc. We will post news about this research on the FrameNet website, including a series of public workshops and presentations on these topics.

References

- Fillmore, C. J. and Baker, C. F. (2010). A frames approach to semantic analysis. In Heine, B. and Narrog, H., editors, *Oxford Handbook of Linguistic Analysis*, pages 313–341. OUP.
- Talmy, L. (2000). *Force Dynamics in Language and Cognition*, volume 1, chapter 7, pages 409–470. MIT Press, Cambridge, MA.