

# Reference Docs Effort

Expanding the *amount of* and *audience for* documentation in the ERG

# ErgSemantics\_Essence

ERS: High-Level Characterization

## Today

ERS does not make explicit lexical semantics beyond what correlates with grammaticized contrasts, e.g. predicate distinctions that co-vary with clear differences in argument frames and types. For example, in the ERS for *The bank with the shortest atm lines is near the river bank*, both instances of *bank* will be represented by `_bank_n_of`.

## A Possible Tomorrow?

Words used in the same grammatical way will produce the same predications. For example, in "The bank with the shortest atm lines is near the river bank", both instances of *bank* will be represented by `_bank_n_of` in the MRS. The ERG only produces different predications when it is clearly a different word based solely on the sentence grammar.

There are many syntactically distinct ways of expressing the same underlying predicate–argument structure, as exemplified in Examples (2a)–(2h) below. So-called diathesis alternations like the dative shift (2b), passivization ((2c) and (2d)), or focus movement ((2e) and (2f)) can lead to stark differences in syntactic structure but have no observable effect on predicate–argument structure...

<Examples>

There are many syntactically distinct ways of expressing the same underlying meaning. This section walks through some examples to show how very different sentences can result in the same MRS document. It isn't important to understand the linguistic terms and reasoning outlined below to see that this kind of abstraction can be very helpful in simplifying the logic used to process the meaning of various sentences.

<Examples>

*Linguistics Background: So-called diathesis alternations like the dative shift (2b), passivization ((2c) and (2d)), or focus movement ((2e) and (2f)) can lead to stark differences in syntactic structure but have no observable effect on predicate–argument structure...*

# Reference Docs

- Introduction to the proposed effort
  - Proposed Goal
  - Proposed Audience
  - Proposed Content
  - Proposed Process
- Issues/Discussion
- Next Steps

# Proposed Goal

- Increase ERG usage by providing documentation that allows a broader set of users to *understand* and *use* the MRS output of the ERG
  - Broad community (non-linguists and linguists)
  - Reduce reliance on limited # of subject matter experts
  - Focus on *practical usage* of the ERG (Python + ACE + Grammar, grammar changes)

# Proposed Audience

- Anyone who wants to build a system using the ERG
  - Research
  - Production
  - Etc.
- Must have computer science experience
  - Documentation is focused on building a system
  - Users capable of finding, installing and using the Delph-In Python library
  - Users who could learn “challenging” languages (Haskell, Prolog, etc)
  - Must have logic 101?

# Proposed Content

- Conceptual/Overview material
  - Variant of: [ErgSemantics](#)
  - With linguistics background (what is an event?)
- Case studies
  - How does one actually use the ERG?
  - Description of the innards of working systems
- Inventory
  - The public interface: the MRS output
  - Variant of: [ErgSemantics Inventory](#)
  - Look at the forum for gaps
- Grammar
  - Nuts and bolts of how to edit the grammar
  - Document syntax for adding custom rules
  - ODan's plate, with reviewers
- Extension Points
  - Areas ripe for enhancement by users
  - Parse ranking, lexical filtering, etc
- Tools and Processes
  - How to use ACE?
  - How to get a build?
  - Who owns what in Delph-in-land?

# Proposed Process Goals

- “Steering committee” that ensures:
  - We are making progress
  - We are going in the right direction
  - Monthly meetings for a while?
  - Members: Olga Zamaraeva, Francis Bond, Dan Flickinger, Luis Morgado da Costa, Alexandre Rademaker, Eric Zinda
- Build a "Sustainable System" for documentation
  - Simple and low overhead: Good enough
  - Allows for broader contributions
  - Recognizes most information is held by a small group of experts
  - Can continue beyond this point-in-time effort



# Proposed Process Details

1. Create a list of topics to write on a public page
  - The list of "todos"
  - Documentation of the process, templates, and process for taking one on
2. Allow contributors to sign up to write a first pass on each
  - Focus on good enough!
  - Recruit a core group of experts
  - ...but anyone in the community can be a contributor
3. Author (or others!) can mark a topic for review
  - "Needs technical review" and/or "Needs accessibility review"
  - Accessibility (i.e. understandability by someone not practiced in the art of linguistics)
  - Allows "editors" to search for docs that need a review pass
  - Mark the doc as reviewed so future readers know who looked at it

# Issues

- How do we think about the Wiki? Best current docs? Archive?
  - Issues: can we erase existing stuff? What if you want to update stuff? Is it OK?
- Where do the docs live?
  - If reference docs are integrated/updated via tools, is wiki the right form?
- What form are the docs written in?
  - For reference docs: comments in source? External markup? Both?
- How do you search?
  - Github search is too simple. More semantic search?
- Who can help write content?
- Are there two sets of some docs for linguists and non-linguists?
  - Can we combine and serve both?
  - Perhaps take a “Tutorial” and “Reference Doc” approach to the conceptual topics?
- How do they stay up to date?
  - Can tools be used to flag out of date reference docs?
  - Are there different sections for reference docs for different grammars?
- Others?