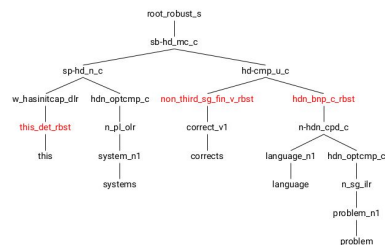


# DELPH-IN Demos



TOP INDEX	$h0$ $e2$
RELS	$\left\langle \begin{matrix} \text{pron}(0:2) \\ \text{LBL} \\ \text{ARG0} \end{matrix} \begin{matrix} h4 \\ x3 \end{matrix} \right\rangle \left\langle \begin{matrix} \text{pronoun}_q(0:2) \\ \text{LBL} \\ \text{ARG0} \\ \text{RSTR} \\ \text{BODY} \end{matrix} \begin{matrix} h5 \\ x3 \\ h6 \\ h7 \end{matrix} \right\rangle \left\langle \begin{matrix} \text{跑}_v \text{步}_\text{sep}_1(3:4) \\ \text{LBL} \\ \text{ARG0} \\ \text{ARG1} \end{matrix} \begin{matrix} h1 \\ e2 \\ x3 \end{matrix} \right\rangle$
HCONS	$\left\langle \begin{matrix} \text{geq} \\ \text{HARG} \\ \text{LARG} \end{matrix} \begin{matrix} h0 \\ h1 \end{matrix} \right\rangle \left\langle \begin{matrix} \text{geq} \\ \text{HARG} \\ \text{LARG} \end{matrix} \begin{matrix} h6 \\ h4 \end{matrix} \right\rangle$



Luis Morgado da Costa  
VU Amsterdam

7<sup>th</sup> July, 2025, Amsterdam

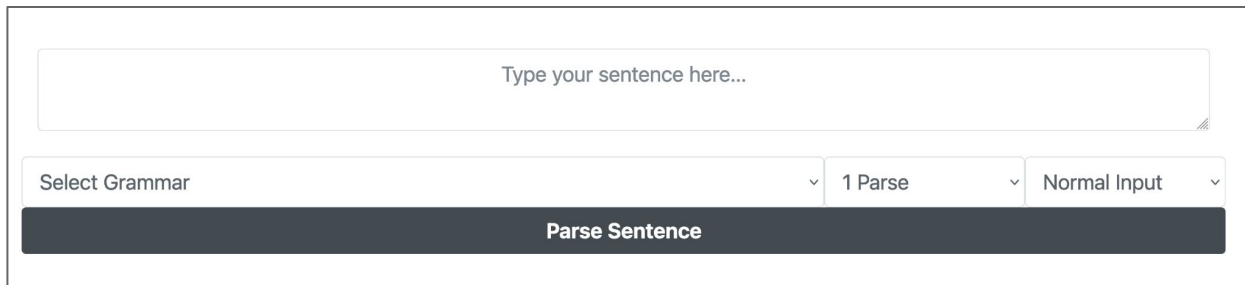


VRIJE  
UNIVERSITEIT  
AMSTERDAM

**VU as a demo:**

**<https://compling.cltl.labs.vu.nl/itell>**

# Grammar/Parser demo



The screenshot shows a web interface for a grammar/parser demo. It features a large text input field at the top with the placeholder text "Type your sentence here...". Below this input field are three dropdown menus: "Select Grammar", "1 Parse", and "Normal Input". At the bottom of the interface is a dark grey button labeled "Parse Sentence".

Here is a list of grammars currently included:

- English Resource Grammar (aka ERG)
- ZHONG (for Mandarin Chinese)
- JACY (for Japanese)
- INDRA (for Indonesian)
- PorGram (for Portuguese)
- YueGram (for Cantonese)
- Spanish Resource Grammar (partial support, through YY input)

I'm happy to add any other grammar that has ACE support.

# DELPH-IN Grammarium

## Grammarium:

- Part of larger web-app (iTELL: GED, CALLIG, etc.)
- Web-based toolkit for grammar development
- PyDelphin, DELPH-IN Viz libraries, ACE
- Grammar Update/Compile (from GitHub)
- YY-input support
- Inspect Profiles/Basic Regression Testing (currently a bit broken with updates)

## Today's Focus:

- Updating grammars through commit branches (**Demo**)
- Talking about metadata (what and where)
- Discuss more broadly what kind of demos we want/need

*I'm sorry Emily  
that it is taking a  
bit too long!*

# Metadata – Is this enough? New/Existing file?

```
GRAMMARS_CONFIG["zhong"]["repo_name"] = "zhong"
GRAMMARS_CONFIG["zhong"]["github_url"] = "https://github.com/delph-in/zhong"
GRAMMARS_CONFIG["zhong"]["grammar_file"] = "zhong.dat"
GRAMMARS_CONFIG["zhong"]["grammar_title"] = "ZHONG"
GRAMMARS_CONFIG["zhong"]["language"] = "Mandarin Chinese"
GRAMMARS_CONFIG["zhong"]["path_to_ace_config"] = "cmn/zhs/ace/config.tdl"
GRAMMARS_CONFIG["zhong"]["path_to_skeletons"] = ["cmn/zhs/tsdb/skeletons"]
GRAMMARS_CONFIG["zhong"]["description"] = ("ZHONG is a mid-sized grammar, with targeted coverage. Much of its "
    "development has been centered around educational data, which means it is "
    "able to provide adequate analyses for most of the sentence patterns "
    "found in early to intermediate educational materials. From a linguistic "
    "point of view, ZHONG offers in-depth analyses of interesting and "
    "well-studied linguistic phenomena including: serial-verbs, 的 (DE) "
    "constructions, 把 (BA) constructions, interactions between aspect and "
    "negation, sentence final particles, and interrogatives.")
GRAMMARS_CONFIG["zhong"]["input_notes"] = ("ZHONG expects its input to be segmented at the word level. This means that "
    "words must be separated by spaces (see some of the examples below). Some "
    "of our analyses expect input to be tokenized differently than most "
    "Mandarin Chinese word segmentation tools (e.g., 这个 should be tokenized "
    "as two tokens 这 and 个). To the best of its ability, the grammar uses "
    "REPP (internal preprocessing) to help further segment the input (e.g., "
    "if 这个 is provided as a single token, it will be internally split as "
    "two tokens).")
GRAMMARS_CONFIG["zhong"]["examples"] = ["我正在看电视呢。",
    "这个商场很大。",
    "我们还要三碗汤。",
    "你会不会开车?",
    "语言学院有多少个系?",
    ]
```

```
GRAMMARS_CONFIG["erg-github"]["repo_name"] = "erg-github"
GRAMMARS_CONFIG["erg-github"]["github_url"] = "https://github.com/delph-in/erg"
GRAMMARS_CONFIG["erg-github"]["grammar_file"] = "erg-latest.dat"
GRAMMARS_CONFIG["erg-github"]["grammar_title"] = "ERG 2023"
GRAMMARS_CONFIG["erg-github"]["language"] = "English"
GRAMMARS_CONFIG["erg-github"]["path_to_ace_config"] = "ace/config.tdl"
GRAMMARS_CONFIG["erg-github"]["path_to_skeletons"] = ["tsdb/skeletons"]
GRAMMARS_CONFIG["erg-github"]["description"] = "This is the latest stabled version of the English Resource Grammar."
GRAMMARS_CONFIG["erg-github"]["input_notes"] = "No special input notes."
GRAMMARS_CONFIG["erg-github"]["examples"] = ["This is a test sentence."]
```

# The metadata feeds the demo

**Summary:** ZHONG is a mid-sized grammar, with targeted coverage. Much of its development has been centered around educational data, which means it is able to provide adequate analyses for most of the sentence patterns found in early to intermediate educational materials. From a linguistic point of view, ZHONG offers in-depth analyses of interesting and well-studied linguistic phenomena including: serial-verbs, 的 (DE) constructions, 把 (BA) constructions, interactions between aspect and negation, sentence final particles, and interrogatives.

**Input:** ZHONG expects its input to be segmented at the word level. This means that words must be separated by spaces (see some of the examples below). Some of our analyses expect input to be tokenized differently than most Mandarin Chinese word segmentation tools (e.g., 这个 should be tokenized as two tokens 这 and 个). To the best of its ability, the grammar uses REPP (internal preprocessing) to help further segment the input (e.g., if 这个 is provided as a single token, it will be internally split as two tokens).

## Examples:

我 正 在 看 电 视 呢 。

这 个 商 场 很 大 。

我 们 还 要 三 碗 汤 。

你 会 不 会 开 车 ？

语 言 学 院 有 多 少 个 系 ？

***Click and Parse  
Demo***

# DELPH-IN Demos

## Discussion Points:

- What kind of demos do we have/want hosted – and where?
  - I heard from Dan that there is some wish in hosting LTDB at the VU (why?)
  - Are there old demos/tools that would be useful + ready to be hosted? [treebank search]
  - The VU is prepared to host tools (within reason...)
- I still have a long wish list from multiple people (Emily, Dan, Mike, Francis, Alex, ...)
  - Dan will probably get some of this checked off in the next couple weeks
  - Unofficially, I am working towards Emily's requests first (because they align nicely with mine – making this tool grammar development friendly)
  - If there are new requests/ideas/suggestions... now would be a nice time
- Maybe we can also dream out loud...
  - What kind of demo/online tool would best help DELPH-IN?

# Thank you!

**We're now open for discussion...**