Language Technology: Research and Development

Dissemination of Research Results

Joakim Nivre

Uppsala University
Department of Linguistics and Philology
joakim.nivre@lingfil.uu.se
Dissemination of Research Results

▶ Why?
▶ Submit results for critical review
▶ Inform other researchers, users, society
▶ Satisfy requirements from funders or customers
▶ Promote research career – publish or perish

▶ To whom?
▶ Other researchers
▶ Potential users
▶ Students
▶ The general public
▶ Funding bodies
▶ Customers
The Receiver

Specific ← Expert → General

Specific ← Novice → General
The Receiver

- Expert
  - Researchers
- Specific
  - Novice
- General

Categories:
- Specific
- General
- Novice
- Expert
- Researchers
The Receiver

- Expert
- Researchers
- Students
- General
- Novice

Specific → General

Language Technology: Research and Development
The Receiver

- Expert
- General
- Researchers
- Students
- General public
- Novice

Specific → General
Novice ← Expert
The Receiver

- Expert
- Researchers
- Funders
- Students
- General public
- Novice
- General

Specific

Language Technology: Research and Development
The Receiver

- Expert
- Researchers
- Students
- General public
- General
- Specific
- Funders
- Customers
- Novice
The Receiver

- Expert
- Research paper
- Researchers
- Specific
- Students
- Customers
- General
- Novice
- General public

Language Technology: Research and Development
The Receiver

- Expert Research paper
- Researchers
- Students
- Customers
- Funders
- Novice
- General public
- Popular science
- General
- Specific

Language Technology: Research and Development
The Receiver

- Novice
- General
- Expert

- Specific
- General

- Project report
- Research paper

- Funders
- Researchers

- Customers

- Students

- Popular science
- General public

- Novice
The Receiver

- Specific
  - Funding
  - Customers

- General
  - Popular science
  - General public

- Expert
  - Research paper
  - Researchers
  - Master’s thesis?
  - Students

- Novice
The Receiver

- Expert
- Research paper
- Researchers
- Master’s thesis
- Students
- Popular science
- General
- General public
- Novice
- Specific
- Project report
- Funders
- Customers
- Novice

Language Technology: Research and Development
How?

Written:
1. Publications (indexed and archived)
2. Internal reports (public or confidential)
3. Digital archives, web pages, etc.

Oral:
1. Lectures (especially at conferences)
2. Demonstrations, posters, discussions, etc.
3. Internal meetings (seminars, workshops)
Written Genres – Single Topic

Papers (short)

1. Journal article – refereed and approved by editorial board
2. Conference paper – often but not always refereed
3. Technical report – usually not refereed

Monographs (long)

1. Book – standards of refereeing depends on publisher
2. Thesis – refereed in examination, may or may not be published
Written Genres – Other

Collections
1. Conference proceedings – collection of conference papers
2. Edited volume – book with different chapter authors

Meta-genres
1. Survey or handbook article
2. Review in scientific journal
3. Bibliography
4. Abstract
Oral Genres

Lecture
▶ Presentation by 1 person followed by discussion (large group)
   1. Conference talk (15–30 min)
   2. Invited talk (45–90 min)

Seminar
▶ Presentation or introduction by 1 or more persons with more or less continuous discussion (small group)

Panel
▶ Short presentations on a set topic from a selected group of persons with questions and opinions from the audience
Mixed Genres

Poster
- Written presentation displayed on poster board
- Oral interaction with interested audience
- Sometimes combined with short talk (1–5 min)

Demonstration
- System demonstration (or similar)
- Oral interaction with interested audience
- Sometimes combined with poster
Requirements on Scientific Reports

▶ Ethics:
  ▶ Sensitive information requires permission and anonymization

▶ Accessibility:
  ▶ Reports should be understandable by target audience

▶ Novelty and relevance:
  ▶ Results should be novel, original, unpublished
  ▶ Relevance to research area should be made clear

▶ Quality:
  ▶ Claims clearly stated and possible to challenge (falsifiability)
  ▶ Claims supported by arguments and/or evidence (justification)
  ▶ Claims not misleading (e.g., by withholding information)
Scientific Writing

Writing takes time (to learn)

- Practice makes perfect – write a lot!
- Writing requires rewriting – start early!

Scientific writing is a standardized genre

- Collect good examples – and study them!
- Copy structure and formulations – but not content!
The Structure of Scientific Publications
The Structure of Scientific Publications

Pre-matter: Title page (abstract, preface, contents)

Post-matter: References (appendices, indexes)
The Structure of Scientific Publications

**Pre-matter:** Title page (abstract, preface, contents)

**Introduction:**
What is the problem/question?
Why is it relevant/interesting?

**Conclusion:**
What is the solution/answer?
Where do we go from here?

**Post-matter:** References (appendices, indexes)
The Structure of Scientific Publications

**Pre-matter:** Title page (abstract, preface, contents)

**Introduction:**
What is the problem/question?
Why is it relevant/interesting?

What has been done before?

**Body:**
How is the problem tackled?
What are the results?

**Conclusion:**
What is the solution/answer?
Where do we go from here?

**Post-matter:** References (appendices, indexes)
The Main Theme

The research question
- is stated in the introduction
- is related to previous research
- motivates the approach taken
- determines the selection of results
- is revisited in the conclusion
The Anatomy of a TACL Style Article

### Title page: title, authors, affiliations

**Token and Type Constraints for Cross-Lingual Part-of-Speech Tagging**

Oscar Ticháček\(^{1}\)\(^{1}\), Dipanjan Das\(^{1}\), Slav Petrov\(^{1}\), Ryan McDonald\(^{1}\), Joakim Nivre\(^{1}\)

\(^{1}\)Swedish Institute of Computer Science

Abstract

We consider the construction of part-of-speech taggers for resource-poor languages. Recently, manually constructed tag dictionaries have been used as type constraints to overcome the scarcity of annotated data in this setting. Here, we show that additional type constraints can be projected from a resource-rich source language to a resource-poor target language via word-aligned blists. We present several models to this end, in particular a partially observed conditional random field model, where aligned tokens are treated as additional features. The resulting taggers achieve a 2% absolute error reduction on the prior state of the art.

### Abstract: self-contained summary

Main text in numbered sections

---

Language Technology: Research and Development 13(21)
The Anatomy of a TACL Style Article

Main text in numbered sections

Acknowledgments (optional)

References (alphabetical by last name)
The Anatomy of a TACL Style Article

Introduction
- State the research problem and relate it to previous research
- Give a synopsis of the rest of the article

Related work
- Model 1: After introduction, before contributions
- Model 2: After contributions, before conclusion

Contributions
- Theory → Method → Results → Discussion

Conclusion
- Evaluate contributions, point to new research directions
References

- Language technology mostly uses the Harvard system
  - Author-year citations in text
  - Alphabetical list of references at the end (no footnotes)
- Citations in the text:
  - Parenthetical: Parsing is hard (Anderson, 2010).
  - Syntactic: Anderson (2010) claims that parsing is hard.
  - More than two authors:
    - In text, use et al.
      Parsing is hard (Anderson et al., 2010).
      Anderson et al. (2010) claims that parsing is hard.
    - All authors in reference list
      Parsing is hard. . . .
Reference List

- Reference list including all (and only) works cited in the text:
  - Journal article: author, year, title, journal, volume, pages
  - Conference paper: author, year, title, proceedings, pages
  - Book chapter: author, year, title, book, publisher, pages
  - Book: author, year, title, publisher
  - Technical report: author, year, title, organization
  - Thesis: author, year, title, type of thesis, school

- Important: Be consistent!
Giving Oral Presentations

Preparation is the key

▶ Think through what you want to say
▶ Formulate key passages in concrete sentences
▶ Prepare audiovisual aids (if relevant)

Practice makes perfect

▶ Rehearse the presentation (many times)
▶ Time the presentation and note any disfluencies
▶ Modify and rehearse until fluent
The Structure of Oral Presentations

Oral presentations are basically structured as written reports but

▶ typically contain less material due to time constraints (especially the background part)
▶ are often less formal and detailed due to real-time processing (the big picture instead of the formal details)
▶ can be more repetitive due to memory limitations (get the take-home message across)

The discussion part:

▶ Listen to the question
▶ Answer the question – if you can
Audiovisual Aids

Slides provide support for the presentation

- Key points and important concepts
- Graphical illustrations (and sound if relevant)
- Material that is hard to present orally (equations, examples)

But remember

- Not too much information (or too small fontsize) on one slide
- Not running text (to be read aloud)
- Slides should support presentation, not vice versa
Geoff Pullum’s Golden Rules

- Don’t ever begin with an apology
- Don’t ever underestimate the audience’s intelligence
- Respect the time limits
- Don’t survey the whole damn field
- Remember that you’re an advocate, not the defendant
- Expect questions that will floor you