# Pair Annotation: Adaption of Pair Programming to Corpus Annotation

**Işın Demirşahin, İhsan Yalçınkaya and Deniz Zeyrek**

**Middle East Technical University, Ankara, Turkey**

disin@metu.edu.tr, yalcinkaya.ihsan@gmail.com, dezeyrek@metu.edu.tr

# Turkish Discourse Bank (TDB)

- 500000-word-subcorpus of Metu Turkish Corpus (Say et al., 2002)

- Annotated for discourse relations in the style of Penn Discourse Tree Bank (Prasad et al., 2008)

  - Discourse connectives

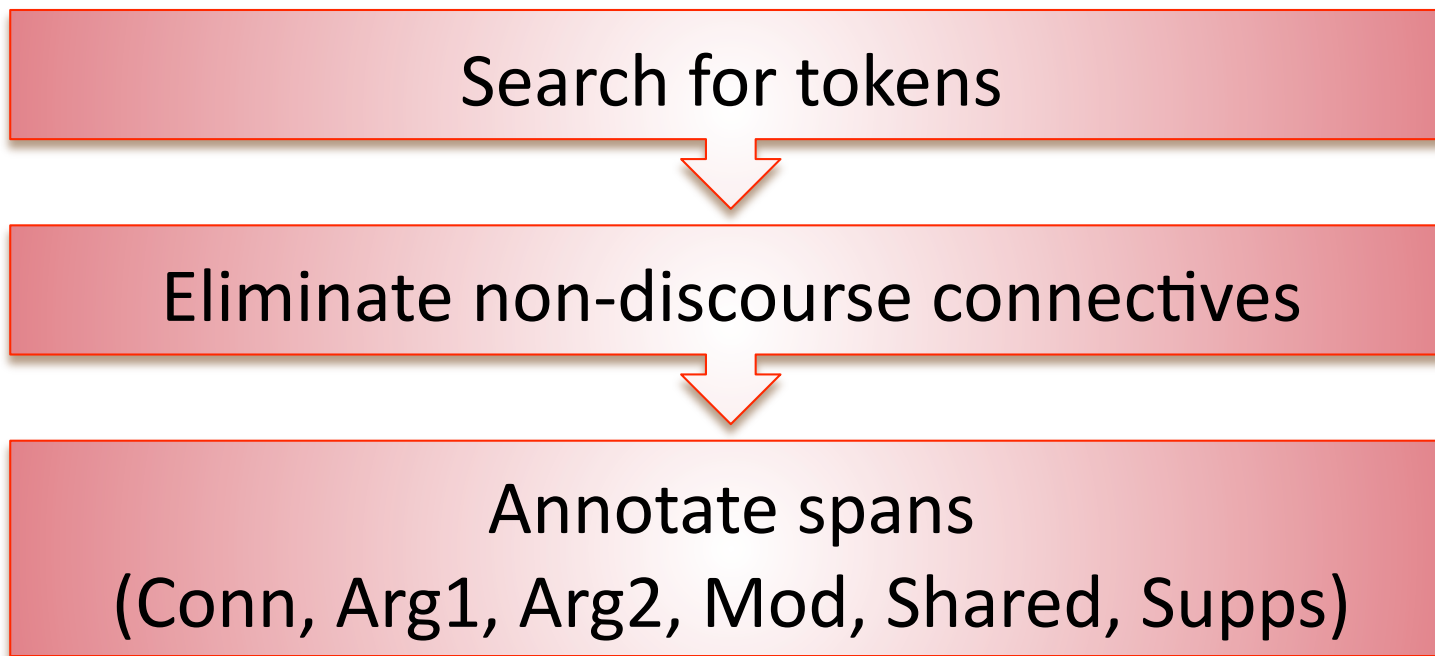  - Arguments -> Abstract objects (Asher, 1993)

# Turkish Discourse Bank (TDB)

- Discourse Annotation Tool for Turkish (Aktaş, et al., 2010)

- Version 1.0:
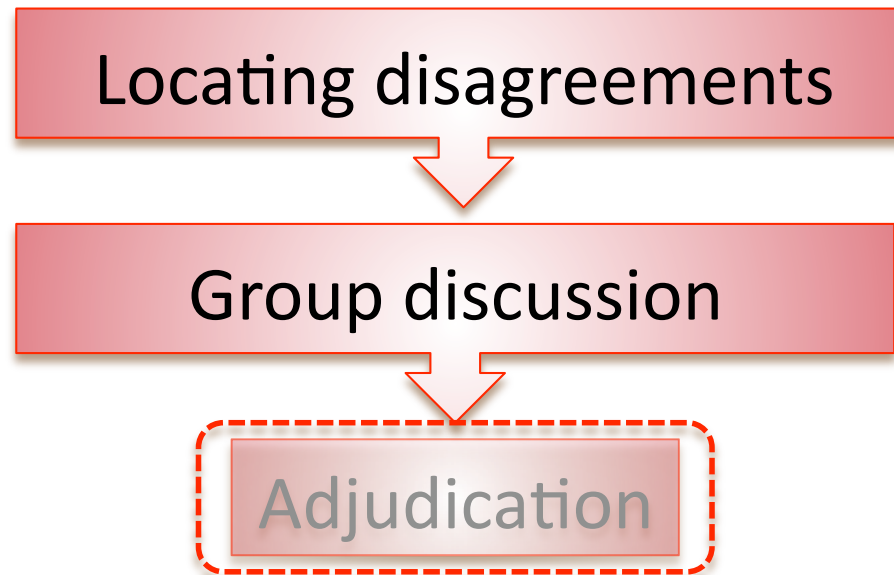  - 147 connectives, 8482 relations
  - http://medid.ii.metu.edu.tr

# Annotation Procedure in the TDB

- Independent Annotation

| Search for tokens |
|---|

↓

| Eliminate non-discourse connectives |
|---|

↓

| Annotate spans<br>(Conn, Arg1, Arg2, Mod, Shared, Supps) |
|---|

# Initial Annotation Procedure in the TDB

- Agreement procedure



6th Linguistic Annotation Workshop

# Initial Annotation Procedure in the TDB



Annotation guidelines

↓

Independent annotations

↓

Agreement procedure

↓

Modification of guidelines

↓

Proof procedure

↓

Gold Standard
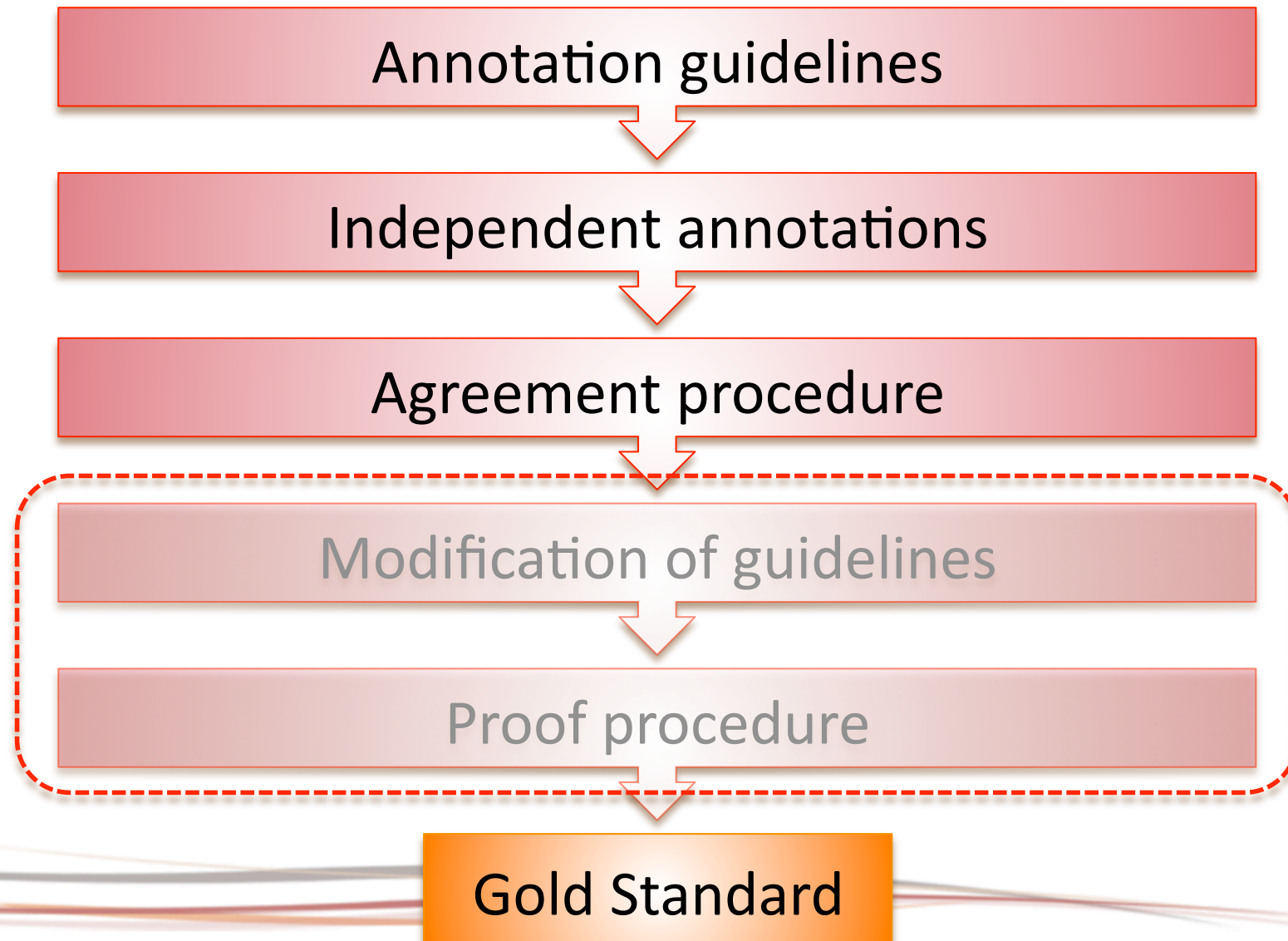
# Common Divergences among Independent Annotations

- Physical errors
  - Word boundaries
  - Spaces and punctuation
  - Pairs of punctuation marks

*"Gidelim" dedim,* <u>ama</u> **dışarıdan yeni sesler geldi.**

*"Let's go"* I said, <u>but</u> **new voices came from outside.**

**Başka kimse olmadığın**<u>dan</u> *iki kadının da yüzü açıktı.*

<u>Since</u> **there was no one else**, *the faces of both women were unveiled*.

# Common Divergences among Independent Annotations

- Ambiguous cases
  - Same relation, with some differences
  - Different relations

| **Salt gülmek** <u>için</u> *gelmişlerdi*. | [Salt] gülmek <u>için</u> ***gelmişlerdi***. |
|---|---|
| *They came* <u>to</u> **just laugh**. | *They came* [only] <u>to</u> **laugh**. |

| *Usumi için ilk tören, Türkiye Gazeteciler Cemiyeti (TGC) önünde düzenlendi.* TGC Başkanı Orhan Erinç, konuşmasında Usumi'nin yokluğunu hissedeceklerini vurguladı. **Usumi'nin cenazesi** [daha] <u>sonra</u> **Sultanahmet Camii'ne götürüldü.** |
|---|
| *The first ceremony for Usumi was arranged in front of the Association of the Journalists of Turkey (TGC).* Orhan Erinç, the chairman the TGC, emphasized that Usumi would be missed. <u>Then</u>, **the Usumi's funeral was moved to the Sultan Ahmed Mosque**. |

# Common Divergences among Independent Annotations

- Missing/invalid annotations
  - Simple overlook
  - Disagreement about the discourse function

| |
|---|
| Yarın demedim, ancak hafta başına hazır olur. |
| I didn't say it would be tomorrow, it won't be ready until the first day of the week. |
| I didn't say it would be tomorrow, but it will be ready on the first day of the week. |

Not D-Conn

D-Conn

| |
|---|
| Kızın saçları *siyah* <u>ve</u> **kıvırcıktı.** |
| The girl's hair was *black* <u>and</u> **curly**. |

# Pair Programming

- Two programmers work as one (Williams, et al, 2000)
  - Driver
  - Navigator
- Joint ownership
  - Credit
  - Responsibility
- The roles are switched periodically

6th Linguistic Annotation Workshop

# Pair Programming

| PROS | CONS |
|---|---|
| More efficient use of coding hours | Increased programming hours (15% - 60%) |
| Increased performance (Pair Jelling) | |
| Decreased communication time between units | |
| Faster production | Increased man-hours |

# Pair Annotation and Pair Programming

- Similarities
  - The driver and the navigator
  - Switching roles
- Difference: Mixed approach
  - One independent annotator
  - One pair of annotators

# Observed Benefits of Pair Annotation

- Higher inter-annotator agreement
  - Less physical errors

- Faster annotation of ambiguous cases
  - Increased discussion during annotation -> decreased discussion during agreement
  - Utilization of *notes* field

- Increased motivation

# Agreement Statistics

- Independent annotation vs. Pair annotation

| Connective | IA – Arg1 | PA – Arg1 | IA – Arg2 | PA – Arg2 |
|---|---|---|---|---|
| *ama* 'but' | 0.832 | 0.956 | 0.901 | 0.969 |
| *sonra* 'after, later' | 0.820 | 0.889 | 0.902 | 0.953 |
| *ve* 'and' | 0.692 | 0.945 | 0.791 | 0.964 |
| *ya da* 'or' | 0.843 | 0.939 | 0.974 | 0.973 |

- Pair-wise averaged K for 3 individual annotators in IA – individual against individual VS K for pair vs. individual in PA

# Agreement Statistics

- Individual annotator vs Pair annotator

| Connective | Ind – Arg1 | Pair – Arg1 | Ind – Arg2 | Pair – Arg2 |
|---|---|---|---|---|
| *aslında* 'actually' | 0.766 | 0.937 | 0.889 | 0.984 |
| *halde* 'despite' | 0.834 | 0.973 | 0.898 | 1.000 |
| *nedeniyle* 'because of' | 0.905 | 0.937 | 0.984 | 0.984 |
| *nedenle* 'for this reason' | 0.952 | 1.000 | 0.987 | 1.000 |
| *ötürü* 'due to' | 1.000 | 1.000 | 0.907 | 0.953 |
| *yüzden* 'because of this' | 0.916 | 0.992 | 0.983 | 1.000 |

- Individual annotator vs. gold standard K VS Pair annotator vs. gold standard K

# Issues in Pair Annotation

- Less sets of independent annotations
- Concern about navigator's contributions
  - Switching roles
  - Periodical feedback
- Biased annotations
  - Keeping record of initial intuitions and disagreement within pair
- Biased agreement process
  - Treating the pair as a single entity
- All solutions are project specific

# Pair Annotation

| PROS | CONS |
|------|------|
| Significant increase of inter-annotator agreements | Possibility that the pair's annotation could be biased |
| Significant increase of annotator vs gold standard agreements | Possibility that the pair could dominate the agreement process |
| Faster agreement process | One less set of annotations |

# Future Work

- Quantitative evaluation of the increase in annotation hours

- Inter-annotator agreement with different pairs

- Inter-annotator agreement between two pairs of annotators.

# Acknowledgements

- We gratefully acknowledge the support of the Turkish Scientific and Technological Research Council of Turkey (TUBITAK) and METU Scientific Research Fund (no. BAP-07-04-2011-005) while developing the Turkish Discourse Bank.

- We thank the Asian Federation of Natural Language Processing (AFNLP) and the National Science Foundation (NSF) for supporting Işın Demirşahin with student grants for travel, registration and accommodation.

# References

- Nicholas Asher. 1993. *Reference to Abstract Objects in Discourse.* Kluwer Academic Publishers.
- Berfin Aktaş, Cem Bozşahin, Deniz Zeyrek. 2010. Discourse Relation Configurations in Turkish and an Annotation Environment. *LAW IV - The Fourth Linguistic Annotation Workshop.* Uppsala, Sweden, July 2010.
- Kent Back. 2000. *Extreme Programming Explained: Embrass Change*. Addison Wesley Longman, Reading, Mass.
- Frederick, P. J. Brooks. 1975. *The Mythical Man-Month*. Addison-Wesley, Reading, Mass.
- Tom DeMarco and Timothy Lister. 1977. *Peopleware*. Dorset House, New York.
- Eleni Miltsakaki, Rashmi Prasad, Aravind Joshi, Bonnie Webber. 2004. Annotating Discourse Connectives and Their Arguments. *HLT/NAACL Workshop on Frontiers in Corpus Annotation*. Boston, MA. May 2004.
- Beata Beigman Klebanov and Eyal Beigman. 2008. From Annotator Agreement to Noise Models. *Computational Linguistics*, 34(3):495–503.
- Rashmi Prasad, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, Bonnie Webber. 2008. The Penn Discourse Treebank 2.0. *LREC'08 - The sixth international conference on Language Resources and Evaluation.* Marrakech, Morocco, May 2008.
- Dennis Reidsma and Jean Carletta. 2008. Reliability measurement without limit. *Computational Linguistics*, 34(3):319–326.
- Bilge Say and Deniz Zeyrek and Kemal Oflazer and Umut Ozge. 2002. Development of a Corpus and a Treebank for Present-day Written Turkish. *ICTL 2002 - 11th International Conference on Turkish Linguistics.* Famagusta, TRNC, August 2002.
- Laurie Williams, Robert R. Kessler, Ward Cunningham, Ron Jeffries. 2000. Strengthening the Case for Pair Programming. *IEEE Software*, July/August 2000:19–25.
- Laurie Williams and Robert R. Kessler. 2000. All I Really Need to Know about Pair Programming I Learned In Kindergarten, *Communications of the ACM*, 43(5):108–114.
- Laurie Williams and Robert R. Kessler. 2003. *Pair Programming Illuminated.* Addison Wesley, Reading, Massachusetts.
- Deniz Zeyrek, Işın Demirşahin, Ayışığı Sevdik-Çallı, Hale Ögel Balaban, İhsan Yalçınkaya and Ümit Deniz Turan. 2010. The Annotation Scheme of the Turkish Discourse Bank and an Evaluation of Inconsistent Annotations. *LAW IV - The Fourth Linguistic Annotation Workshop* Uppsala, Sweden, July 2010.

# Additional Reference

- Yalçınkaya, İhsan. 2010. *An Inter-Annotator Agreement Methodology for the Turkish Discourse Bank. Unpublished MS Thesis, Middle East Technical University.*

# Initial Annotation Procedure in the TDB

- One search token -> Multiple connectives

- *Halde*

  - *-DIğI halde* 'although, despite'

  - *o halde* 'then, in that case'

  - *şu halde* 'in the current situation, in this specific case'

  - *aksi halde,* 'otherwise'

  - *bir halde* 'in such a manner that'

# Initial Annotation Procedure in the TDB

**Doğu Beyazıt'da gecelediğimiz** <u>halde</u> *bir dünya şaheseri olan İshak Paşa medresesini göremeden Ankara'ya döndük.*

<u>Although</u> **we spent the night in Doğu Beyazıt**, *we returned to Ankara without seeing the İshak Paşa Medresseh, which is a masterpiece.*

Feyzi Bey *böyle bir durumda mebusluktan istifa edeceğini,* <u>aksi halde</u> [de] **Falih Rıfkı Bey'in istifa etmesi gerektiğini** belirtmiş.

Mr. Fevzi stated that *in such a situation he would resign from parliament membership,* <u>otherwise</u> **Mr. Falih Rıfkı would have to resign**.

O gün akşama kadar ne yapacağımı bilmez bir halde dolaştım evin içinde.

I walked around the house till evening that day, not knowing what to do.

Not annotated!