LREC 2022, 13th Conference on Language Resources and Evaluation, Marseille, France, 20-25 June 2022

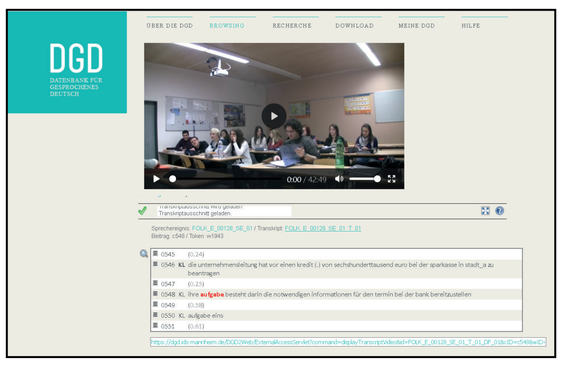
Querying Interaction Structure: Approaches to Overlap in Spoken Language Corpora

Elena Frick[1], Henrike Helmer[1] and Thomas Schmidt[2]

[1] IDS Mannheim (DE), [2] University of Basel (CH)

BACKGROUND

Research Project „ZuMult“

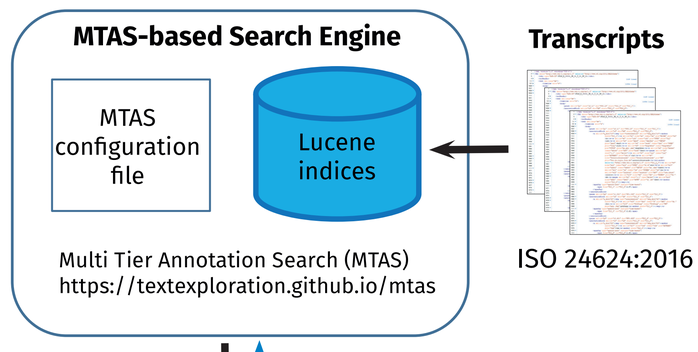


ZuMult stands for German „**Zu**gänge zu **mult**imodalen Korpora gesprochener Sprache“ – „Access to Multimodal Spoken Language Corpora”

Aim of the project:
a backend software architecture for a unified access to spoken language resources (natural interaction) located in different repositories (incl. **search functionality**)

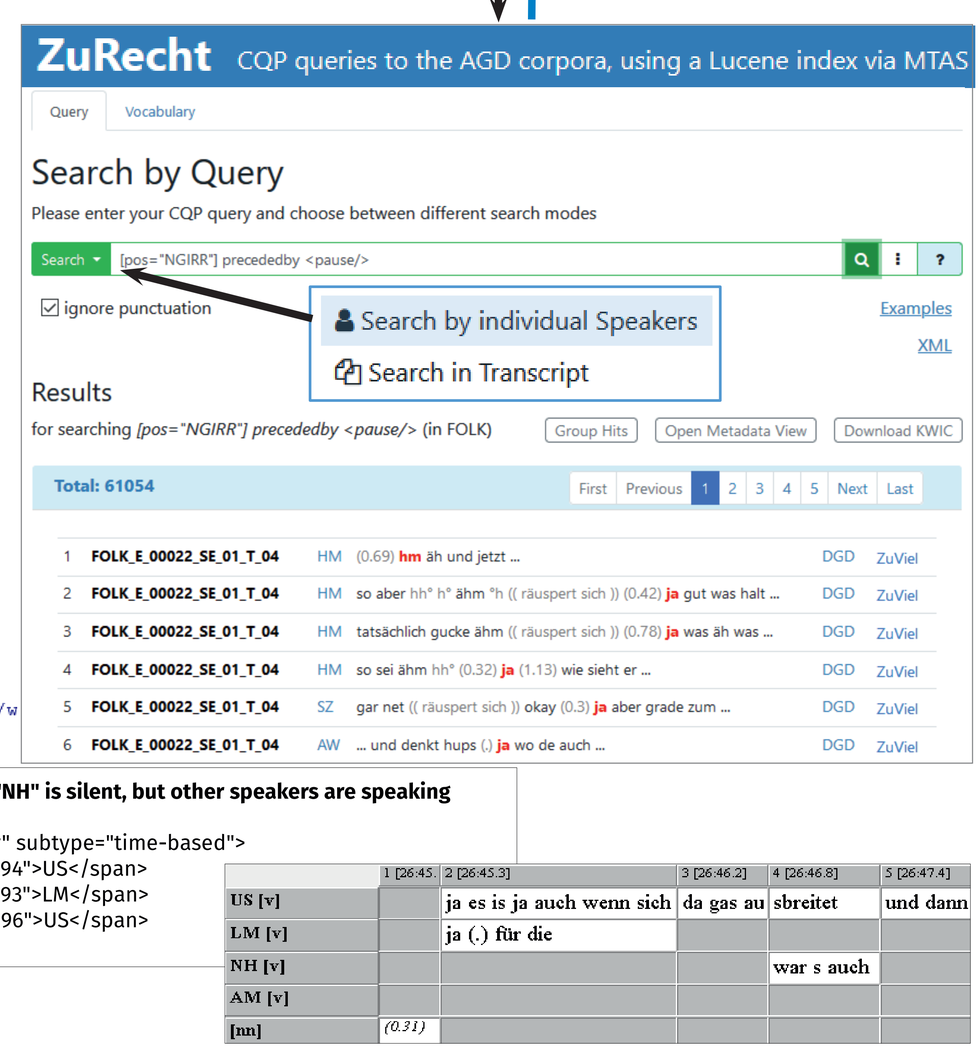
ZuMult-Blog at https://zumult.org/

APPROACH



([norm="oder"] !within <speaker-overlap/>) followedby <para/>{0,5}<another-speaker/>
looks for any transcribed form of German ‘oder’ (engl. ‘or’) occurring outside of an overlap at the last position before speaker change; 'para' stands for <pause>-, <vocal>- and <incident>-elements which can occur in the transcription between two speaker contributions.

(<annotationBlock/> containing ([word=".*" & !pos="(NGIRR|NGHES|XY)"] !within <speaker-overlap/>)) precededby (<another-speaker/><para/>{0,5})
looks for turn-taking by one of the non-speakers whose contribution contains at least one word token that occurs outside of the speaker overlap and is not a non-word, hesitation, interjection or responsive particle.




```
<pause dur="PT0.31S" end="TLI_992" xml:id="p301"/>
<annotationBlock xml:id="c667" end="TLI_996" start="TLI_992" who="US">
   <u>
    <seg>
      <anchor synch="TLI_992" type="ol-start"/>
      <w xml:id="w3007" type="ol-in" lemma="ja" norm="ja" pos="NGIRR">ja</w>
      <w xml:id="w3008" type="ol-in">es</w>
      <w xml:id="w3009" type="ol-in">is</w>
      <w xml:id="w3010" type="ol-in" >ja</w>
      <w xml:id="w3011" type="ol-in" >auch</w>
      <w xml:id="w3012" type="ol-in" >wenn</w>
      <w xml:id="w3013" type="ol-in" >sich</w>
      <anchor synch="TLI_993" type="ol-end"/>
      <w xml:id="w3014">da</w>
      <w xml:id="w3015">gas</w>
      <w xml:id="w3016" type="ol-in">au<anchor synch="TLI_994" type="ol-start"/>sbreitet</w
      <anchor synch="TLI_995" type="ol-in"/>
      <w xml:id="w3017" type="ol-in" >und</w>
      <w xml:id="w3018" type="ol-in">dann</w>
      <anchor synch="TLI_996" type="ol-end"/>
    </seg>
   </u>
   <spanGrp type="speaker-overlap" subtype="time-based">
      <span from="TLI_992" to="TLI_993">LM</span>
      <span from="TLI_994" to="TLI_995">NH</span>
      <span from="TLI_995" to="TLI_996">AM</span>
   </spanGrp>
</annotationBlock>
<annotationBlock end="TLI_993" start="TLI_992" who="LM" xml:id="c668">
   <u>
    <seg>
      <anchor synch="TLI_992" type="ol-start"/>
      <w xml:id="w3019" type="ol-in" lemma="ja" norm="ja" pos="NGIRR">ja</w>
      <pause rend="(.)" type="micro" xml:id="p302"/>
      <w xml:id="w3020" type="ol-in">für</w>
      <w xml:id="w3021" type="ol-in">die</w>
      <anchor synch="TLI_993" type="ol-end"/>
    </seg>
   </u>
   <spanGrp type="speaker-overlap" subtype="time-based">
      <span from="TLI_992" to="TLI_993">US</span>
   </spanGrp>
</annotationBlock>
<annotationBlock xml:id="c669" end="TLI_995" start="TLI_994" who="NH">
  <u>
   <seg>
     <anchor synch="TLI_994" type="ol-start"/>
     <w xml:id="w3022" type="ol-in">war</w>
     <w xml:id="w3023" type="assimilated ol-in" >s</w>
     <w xml:id="w3024" type="ol-in" >auch</w>
     <anchor synch="TLI_995" type="ol-end"/>
   </seg>
  </u>
  <spanGrp type="speaker-overlap" subtype="time-based">
     <span from="TLI_994" to="TLI_995">US</span>
  </spanGrp>
</annotationBlock>
```

Time intervals in which speaker "NH" is silent, but other speakers are speaking

```
<spanGrp type="another-speaker" subtype="time-based">
  <span from="TLI_992" to="TLI_994">US</span>
  <span from="TLI_992" to="TLI_993">LM</span>
  <span from="TLI_995" to="TLI_996">US</span>
<spanGrp/>
```

	1 [26:45.]	2 [26:45.3]	3 [26:46.2]	4 [26:46.8]	5 [26:47.4]
US [v]		ja es is ja auch wenn sich	da gas au	sbreitet	und dann
LM [v]		ja (.) für die			
NH [v]				war s auch	
AM [v]					
[nn]	(0.31)				

	9 [51:57.5]	10 [51:58.]	11 [51:58.5]	12 [51:58.9]	13 [52:00.2]	14 [52:00.5]	15 [52:01.4]
AM [v]			((lacht))			ja des nein nein	es is
US [v]	ja so (.)	nee de	s klingt d	es klingt jetzt vielleicht hart ich weiß	nich wie ich s	besser ausdrücken	sol
NH [v]		((lacht)) das jetz vielleicht über			trieben		ja
[nn]							

	1 [52:41.5]	2 [52:43.	3 [52:43.9]	4 [52:44.4]	5 [52:45.
US [v]	((Lachansatz)) (.) der muss schal schmecken (.) ((Lachansatz))	°h ((lacht))		also trinken ((lacht))	
NH [v]	also der muss (.) eigentlich muss er weg	°h ((lacht))			
AM [v]			war der	war der im kühlschrank also is (.) is der	kalt
[nn]					

Speaker Overlaps: segment-based vs. contribution-based approach

[word.type=".*ol-in.*"]
looks for word tokens within overlaps; the search pattern containing regular expression characters ‘.’ from both sides of ‘ol-in’ is important to match also type-attributes containing multi-word values*

[norm="bitte" & word.type=".*ol-in.*"]
looks for any transcribed form of ‘bitte’ within overlaps

<annotationBlock/> containing [word.type=".*ol-in.*"]
looks for all speaker contributions containing overlaps

<speaker-overlap/>
looks for all spans annotated as speaker overlap

<speaker-overlap/> containing [lemma="(Herr|Frau)"]
looks for all spans annotated as speaker overlap and containing any forms of 'Herr' or 'Frau'

<speaker-overlap>[norm="also"]
looks for any transcribed form of 'also' at the beginning of speaker overlaps

<speaker-overlap="SZ"/>
looks for all token sequences overlapping with the contributions of the speaker ‘SZ’

More about the project:

Fandrych, C., Frick, E., Kaiser, J., Meißner, C., Portmann, A., Schmidt, T., Schwendemann, M., Wallner, F., and Wörner, K. (2022). ZuMult: Neue Zugangswege zu Korpora gesprochener Sprache. In Kämper, H. et al. (Eds.), Sprache in Politik und Gesellschaft: Perspektiven und Zugänge. Jahrbuch des Instituts für Deutsche Sprache 2021. Berlin etc.: de Gruyter.

Frick, E. and Schmidt, T. (2020). Using Full Text Indices for Querying Spoken Language Data. In *Proceedings of the LREC 2020 Workshop, Language Resources and Evaluation Conference, 11–16 May 2020, 8th Workshop on Challenges in the Management of Large Corpora (CMLC-8)*, pages 40–46, Paris: European Language Resources Association (ELRA).

Contact:
Elena Frick M.A.
Abteilung Pragmatik
Leibniz-Institut für Deutsche Sprache
Postfach 10 16 21
68016 Mannheim, Germany

Phone: +49 621 1581-137
Fax: +49 621 1581-200
frick@ids-mannheim.de

Street Address:
Leibniz-Institut für Deutsche Sprache
R 5, 6-13
68161 Mannheim, Germany

Phone: +49 621 1581-0
Fax: +49 621 1581-200
info@ids-mannheim.de
www.ids-mannheim.de

Leibniz Association