

APPENDIX to

## Communicating the unknown

### An interdisciplinary annotation study of uncertainty in the coronavirus pandemic

Marcus Müller<sup>i</sup>, Sabine Bartsch<sup>i</sup>, Jens O. Zinn<sup>ii</sup>

<sup>i</sup>TU Darmstadt | <sup>ii</sup>University of Melbourne

#### Abstract

This paper presents an annotation approach to examine uncertainty in British and German newspaper articles on the coronavirus pandemic. We developed a tagset in an interdisciplinary team from corpus linguistics and sociology. After working out a gold standard on a pilot corpus, we apply the annotation to the entire corpus drawing on an “annotation-by-query” approach in CQPWeb, based on uncertainty constructions that have been extracted from the gold standard data. The annotated data are then evaluated and sociologically contextualised. On this basis, we study the development of uncertainty markers in the period under study and compare media discourses in Germany and the UK. Our findings reflect the different courses of the pandemic in Germany and the UK as well as the different political responses, media traditions and cultural concerns: While markers of fear are more important in British discourse, we see a steadily increasing level of disagreement in German discourse. Other forms of uncertainty such as ‘possibility’ or ‘probability’ are similarly frequent in both discourses.

To appear in

**International Journal of Corpus Linguistics**

DOI: <https://doi.org/10.1075/ijcl>

## Corpus Query Language (CQL) Queries used in the paper

Cf.

Evert, S. & The CWB Development Team (2020). The IMS Corpus Workbench (CWB). CQP Query Language Tutorial. CWB Version. [http://cwb.sourceforge.net/files/CQP\\_Tutorial.pdf](http://cwb.sourceforge.net/files/CQP_Tutorial.pdf)

### Personal\_Anxiety

UK

“( <s>[ ]\*( [lemma="fear|worry|anxi.\*|worried|frighten.\*|nervous.\*|scary|panic"%c&word!="anxiolytic.?|anxietyuk.org"%c]([word!="no"][word="concerns"%c&pos!="NP.?|VVZ"])[word="concerned"&pos="JJ"])\*</s>)( <s>[ ]\*[word="concerned"&pos="V.\*"][word="about|that"])\*</s>” returned 9,479 matches in 4,952 different texts.

Categorised queries:

“( <s>[ ]\*[word="concerned"&pos="V.\*"][word!="about|that"])\*</s>)( <s>[ ]\*[word!="of|no"][word="concern"&pos="NN"])\*</s>” returned 1,162 matches in 987 different texts (in 9,133,404 words [13,379 texts]; frequency: 127.23 instances per million words), manually categorised as “UK\_pers\_anxiety\_concern\_new\_1\_true” (1,099 hits)

“(`<s>[]*[word="concerns"%c&pos="VVZ"][]*</s>`)” returned 22 matches in 22 different texts, manually categorised as “UK\_pers\_anxiety\_concerns\_VVZ\_true” (10 hits)

“(`<s>[]*[lemma="insecur.*"][]*</s>`)” returned 77 matches in 72 different texts (in 9,133,404 words [13,379 texts]; frequency: 8.43 instances per million words), manually categorised as “UK\_pers\_anxiety\_insecurity\_new1\_anxiety” (13 hits)

“(`[lemma="uncertain.*"%c]`)” returned 1,140 matches in 918 different texts (in 9,133,404 words [13,379 texts]; frequency: 124.82 instances per million words), sorted on *node word* (1,140 hits), manually categorised as “UK\_pers\_uncertain\_ty\_anxiety” (82 hits)

## GER

“(`<s>[]*[lemma="befürchten|besorgnis.*|fürchten|Angst|Befürchtung|Furcht|Sorge|Nervosität|Unsicherheit|ängstlich|beängstigend"%c][]*</s>`):: match.text\_id="(FAZ|SP|WE)\_\d+” returned 5,131 matches in 2,901 different texts

“(`<s>[]*[lemma="unsicher"%c][]*</s>`):: match.text\_id="(FAZ|SP|WE)\_\d+” returned 198 matches in 190 different texts (in 17,918,177 words [27,137 texts]; frequency: 11.05 instances per million words), manually categorised as “GER\_sit\_vagueness\_unsicher\_false” (46 hits), manually categorised as “GER\_pers\_anxiety\_unsicher\_true” (17 hits)

## Personal Disagreement

### UK

“(`<s>[]*[lemma="debate.*|disagree.*|conflict|controvers.*"][]*</s>`)(`<s>[]*[lemma="argue.*"][]*[word="against"][]*</s>`)” returned 1,245 matches in 949 different texts

“(`[word="no|neither|nor"%c][]{0,3}[lemma="agreement"]`)” returned 16 matches in 16 different texts, manually categorised as “UK\_disagreement\_no\_agreement\_true” (5 hits)

### GER

“(`<s>[]*[lemma="debattieren|diskutieren|streit.*|Auseinandersetzung|Debatte|Konflikt|kontrovers"%c][]*</s>`):: match.text\_id="(FAZ|SP|WE)\_\d+” returned 3,076 matches in 1,929 different texts (in 17,918,177 words [27,137 texts]; frequency: 171.67 instances per million words)

## Personal Doubt

### UK

“(`<s>[]*([word!="nobody|without|no|little"%c]{4}[lemma="doubt.*"]|[lemma="doubtful|doubter|sceptic.*"%c])[]*</s>`)” returned 595 matches in 528 different texts.

### GER

“(`<s>[word!="kein.*|ohne"]*[lemma=".*zweifeln|Zweifel|skeptisch|Skepsis"&lemma!="zweifellos"%c][]*</s>`):: match.text\_id="(FAZ|SP|WE)\_\d+” returned 1,144 matches in 970 different texts (in 17,918,177 words [27,137 texts]; frequency: 63.85 instances per million words)

## Personal\_Ignorance

UK

“<s>[]\*([word="n't|not|no|none|nobody"%c][[]]{0,2}[lemma="aware|know.\*|understand|idea|evidence|data"])[word="ignorance|unaware|unknown"%c]([word="no|nobody"%c][word="one"]){0,1}[word="can"])[word="imagine|foresee|say"])]\*</s>” returned 2,106 matches in 1,619 different texts

Categorised queries:

“<s>[]\*[lemma="uncertain.\*"%c][[]]\*</s>” returned 1,140 matches in 918 different texts, manually categorised as “UK\_pers\_uncertain\_ty\_ignorance” (460 hits)

GER

“((<s>[]\*[lemma="unbekannt|ungeklärt|unklar"%c][[]]\*</s>)|(<s>[]\*[word="nicht"%c][word="bekannt"])]\*</s>)|(<s>[word!="nicht"]\*[word="erforderlich"%c][[]]\*[lemma="Untersuchung"%c][[]]\*</s>)|(<s>[]\*[lemma="Untersuchung"%c][word!="nicht"]\*[word="erforderlich"%c][[]]\*</s>)|(<s>[]\*[lemma="keine"%c][[]]\*[lemma="Beweis"])]\*</s>)|(<s>[]\*[lemma="sein"])[[]]\*([word="nicht"%c]([word="in"])[word="keiner"])[word="Weise"])]{0,1}[word="erwiesen|sicher|gesichert"])]\*</s>)|(<s>[]\*([word="nicht"])([word="nur"])[word="schwer"])]{0,1}[word="abschätzen"])]\*</s>)|(<s>[]\*[word="sich"])[[]]{0,1}[word="die"])[lemma="Frage"])]\*</s>)|(<s>[]\*[lemma="fragen"%c][word="sich"])]\*</s>)|(<s>[]\*[lemma="keine.\*"%c][[]]{0,3}[lemma="Aussage"])]\*</s>)|(<s>[]\*[lemma="keine|niemand|nicht"%c][[]]{0,3}[lemma="wissen"])]\*</s>)|(<s>[]\*[lemma="wissen"%c][[]]{0,3}[lemma="niemand|nicht"])]\*</s>)|(<s>[]\*[lemma="Ergebnis"%c][[]]{0,7}[lemma="abwarten"])]\*</s>)|(<s>[]\*[lemma="keine"%c][[]]{0,3}[lemma="Angabe|Information|Evidenz"])]\*</s>)) :: match.text\_id="(FAZ|SP|WE)\_\d+” returned 2,563 matches in 1,874 different texts (in 17,918,177 words [27,137 texts]; frequency: 143.04 instances per million words)

## PERSONAL\_Presumption

UK

“<s>[]\*([lemma="apparently|assum.\*|guess.{0,2}|presum.\*|suspicion|supposition|seem"%c&word!="presumptuous"])[lemma="suspect"&pos="V.+"])([lemma="estimat.\*"][word="that"])([word="is|are"][word="believed|thought|supposed"])([word="suppose.\*"%c][word!="to"%c])([lemma="must"&pos="MD"])[word="have"][pos="V.N"])]\*</s>” returned 6,141 matches in 3,802 different texts.

“<s>[]\*[word!="I|you|if"%c][pos!="SENT"]\*[lemma="might"&pos="MD"])]\*</s>” returned 3,599 matches in 2,426 different texts.

Categorised queries:

“<s>[]\*[pos!="PP|NP.\*|TO"]{3}[word="suggest.?"&pos="V.\*"][word="that"])]\*</s>”: this query returned 244 matches in 230 different texts, manually categorised as “UK\_pers\_presumption\_suggest\_true” (236 hits)

GER

“(<s>[]\*[word="offenbar|vermutlich"%c&pos="AD.+"])]\*</s>)|(<s>[]\*[word="davon"])[word="aus|ausgehen"])]\*</s>)|(<s>[]\*[lemma="ahnen|vermuten"])]\*</s>)|(<s>[lemma!="Sonne|Mond"]\*[lemma="scheinen"])[lemma!="Sonne|Mond"]\*</s>)|(<s>[]\*([lemma="hindeuten"])(

[lemma="deuten"][] {0,2} [word="darauf"] [word="hin"] ( ([word="deutet"] [word="auf"] [] {0,7} [word="hin"] ) [] \* </s> ) ( <s> [lemma!="wenn|sagen|meinen"] \* [word="dürfte"] [lemma!="wenn|sagen|meinen"] \* </s> ) ( <s> [] \* [word="müsste.?" %c ] [] {0,2} [word="doch" %c ] [] \* </s> ) ( <s> [] \* [lemma="Ahnung|. \*abschätzung|Schätzung|Spekulation|Vermutung|Verdacht|Verdachtsfall" %c ] [] \* </s> ) ( <s> [] \* [word!="sehr|mich|sich|uns|super|total|vollkommen|absolut|richtig"] [word="wohl" &pos="ADV" ] [] \* </s> ) :: match.text\_id="(FAZ|SP|WE)\_\d+" returned 9,427 matches in 4,789 different texts (in 17,918,177 words [27,137 texts]; frequency: 526.11 instances per million words)

Categorised queries:

<s> [lemma!="wenn|sagen|meinen"] \* [word="dürften"] [lemma!="wenn|sagen|meinen"] \* </s> :: match.text\_id="(FAZ|SP|WE)\_\d+" returned 781 matches in 690 different texts (in 17,918,177 words [27,137 texts]; frequency: 43.59 instances per million words), manually categorised as "GER\_pers\_presumption\_duerften\_true" (722 hits)

<s> [] \* [word!="ganz"] [word="bestimmt" %c &pos="AD.+" ] [] \* </s> :: match.text\_id="(FAZ|SP|WE)\_\d+" returned 87 matches in 86 different texts (in 17,918,177 words [27,137 texts]; frequency: 4.86 instances per million words), manually categorised as "GER\_pers\_presumption\_bestimmt\_true" (72 hits)

[lemma="schätzen|abschätzen"] :: match.text\_id="(FAZ|SP|WE)\_\d+" returned 718 matches in 645 different texts (in 17,918,177 words [27,137 texts]; frequency: 40.07 instances per million words), manually categorised as "GER\_pers\_presumption\_schaetzen\_true" (645 hits)

## SITUATIONAL\_Danger

UK

<s> [] \* [lemma="risk. \*|danger. \*|harmful|hazard. \*|peril. \*|threat.?" %c ] [] \* </s> returned 8,979 matches in 4,805 different texts.

Categorised queries:

<s> [] \* [lemma="uncertain.?" %c ] [] \* </s> returned 1,140 matches in 918 different texts, manually categorised as "UK\_pers\_uncertain\_ty\_danger" (571 hits)

<s> [] \* [lemma="insecur.?" ] [] \* </s> returned 77 matches in 72 different texts, manually categorised as "UK\_pers\_anxiety\_insecurity\_new1\_danger" (64 hits)

GER

<s> [] \* [lemma="gefährden|gefährlich|. \*gefahr|. \*risiko|. \*gefährdung|riskant|riskieren" %c ] [] \* </s> :: match.text\_id="(FAZ|SP|WE)\_\d+" returned 5,704 matches in 3,195 different texts (in 17,918,177 words [27,137 texts]; frequency: 318.34 instances per million words)

<s> [pos!="NE"] \* [lemma="drohen" %c ] [pos!="NE"] \* </s> :: match.text\_id="(FAZ|SP|WE)\_\d+" returned 795 matches in 696 different texts (in 17,918,177 words [27,137 texts]; frequency: 44.37 instances per million words)

( <s> [] \* [lemma="drohen" %c ] [] \* [pos="NE" ] [] \* </s> ) ( <s> [] \* [posge="NE" ] [] \* [lemma="drohen" %c ] [] \* </s> ) :: match.text\_id="(FAZ|SP|WE)\_\d+" returned 496 matches in 448 different

texts (in 17,918,177 words [27,137 texts]; frequency: 27.68 instances per million words), manually categorised as “sit\_danger\_drohen\_true” (448 hits)

“<s>[\*[lemma="unsicher"%c][\*</s>:: match.text\_id="(FAZ|SP|WE)\_\d+” returned 198 matches in 190 different texts (in 17,918,177 words [27,137 texts]; frequency: 11.05 instances per million words), manually categorised as “GER\_sit\_vagueness\_unsicher\_false” (46 hits), manually categorised as “GER\_pers\_danger\_unsicher\_true” (24 hits)

## SITUATIONAL\_Chance

UK

“<s>[\*([word="a|best|good|last|better|your|their|second|real|realistic|greater|possible|unique|great|excellent|wonderful"])[lemma="chance"]][word="promising"&pos="JJ"])[lemma="opportunity"]][\*</s>” returned 1,717 matches in 1,327 different texts

Categorised query:

“<s>[\*([word!="a|best|good|little|last|better|your|their|second|real|realistic|greater|possible|unique|great|excellent|wonderful"])[lemma="chance"%c][\*</s>” returned 913 matches in 796 different texts, manually categorised as “UK\_sit\_opportunity\_chance\_positive” (473 hits)

GER

“(<s>[\*[word!="bei"]][word!="jeder"])[lemma="Gelegenheit"%c][\*</s>)|(<s>[\*[lemma="vielversprechend|.chance"%c][\*</s>] :: match.text\_id="(FAZ|SP|WE)\_\d+” returned 1,599 matches in 1,231 different texts (in 17,918,177 words [27,137 texts]; frequency: 89.24 instances per million words)

“<s>[pos!="NE|PPER"][\*[lemma="versprechen|verheißen"&pos="VVFIN"%c][pos="ART"]]? [pos="ADJA"]? [pos="NN"])[pos!="NE|PPER"][\*</s> :: match.text\_id="(FAZ|SP|WE)\_\d+” returned 43 matches in 40 different texts (in 17,918,177 words [27,137 texts]; frequency: 2.40 instances per million words), manually categorised as “GER\_sit\_Chance\_versprechen\_true” (12 hits)

## SITUATIONAL\_Possibility

UK

(<s>[\*[lemma="eventually|maybe|possibly|potentially|eventuality|possibility|potential"]][\*</s>)|(<s>[word!="said|told|added|announced|confirmed|admitted|agreed|accused|instisted|promised|tweeted|posted|approved|pointed|acknowledged|stressed|responded|replied"])\*([word="could"%c][word!="I|you|we|n't|not"])[\*</s>)|(<s>[\*[word="may"&pos="MD"]][word!="I|you|we"])[\*</s>)

GER

“(<s>[\*[lemma="möglich.{0,2}|möglichlicherweise|Möglichkeit"%c&word!="möglichst"%c][\*</s>)|(<s>[word!="müsse|müssten|sei|habe|meint\*|sagt\*|behauptet\*"%c]\*[word="könnte|können"%c][word!="müsse|müssten|sei|habe|meint\*|sagt\*|behauptet\*"%c]\*</s>)|(<s>[\*[word="kann"]][{0,2}[word="sein"]][\*</s>):: match.text\_id="(FAZ|SP|WE)\_\d+” returned 13,437 matches in 5,689 different texts (in 17,918,177 words [27,137 texts]; frequency: 749.91 instances per million words)

→ matches all lexical possibility markers as well as instances of *können* (Konj. II), except of those appearing in indirect speech (which point mostly to different modal backgrounds)

## SITUATIONAL\_Probability

UK

“<s>[\*[lemma=".\*likel.\*|.\*.probab.\*"]]\*</s>” returned 6,082 matches in 3,831 different texts.

“<s>[\*([word!="a|best|good|little|last|better|your|their|second|real|realistic|greater|possible|unique|great|excellent|wonderful"])[lemma="chance"%c])]\*</s>” returned 913 matches in 796 different texts, manually categorised as “UK\_sit\_chance\_possibility” (338 hits)

GER

<s>[\*[lemma=".\*wahrscheinlich.\*"%c]]\*</s>)(<s>[\*[word="nahezu|praktisch"%c][word="ausgeschlossen"]]\*</s>): match.text\_id="(FAZ|SP|WE)\_\d+” returned 1,391 matches in 1,123 different texts (in 17,918,177 words [27,137 texts]; frequency: 77.63 instances per million words)

## SITUATIONAL\_Vagueness

UK

“<s>[\*((([word="around"][pos="CD"])[word="almost"])([word="at"][word="least"])([word="at"][word="the"]?[word="most"])([word="at"][word="some"][word="point"])([word="in"][word="some"][word="cases"])([word="more"][word="than"])[word="partially|partly|somehow|sometime|somewhere"])([word="some"][word="degree"][word="of"])])\*</s>” returned 12,189 matches in 6,603 different texts.

GER

“(<s>[\*[lemma="beinahe|fast|höchstens|irgendwann|irgendwie|irgendwo|maximal|mindestens|nahezu|teilweise|relativ|vage|Vagheit"%c]]\*</s>)(<s>[\*[word="zum"][word="Teil"]]\*</s>): match.text\_id="(FAZ|SP|WE)\_\d+” returned 8,300 matches in 4,655 different texts (in 17,918,177 words [27,137 texts]; frequency: 463.22 instances per million words)