Universität Münster

SPRACHE UND KONFESSION IM RADIO (DFG)

RELIGION ON AIR

Corpus Linguistic Approaches to the Analysis of 20,000 Radio Sermons

Dr. Anna-Maria Balbach, Germanistisches Institut, Universität Münster
Dr. Jan Oliver Rüdiger, Institut für deutsche Sprache (IDS), Mannheim

1. Background: Christianity has strongly influenced the German language for centuries. The Reformation and the establishment of different denominations have also led to differences in language use between Catholics and Protestants. The relationship between language and denomination has not yet been studied for the present-day language. The aim of this project is to analyze texts of today's public radio sermons at different linguistic levels and to find out if denominational differences in language use can still be detected today – 500 years after the Reformation.

2. Research Questions: To identify denominational linguistic differences, we examine the radio sermons by asking the following questions: How do the Catholic and Protestant authors structure their sermons? What content do they consider relevant and how is it conveyed linguistically? Do the denominations use the same content and linguistic means, or do denominational preferences emerge?

3. Research Corpus: In our project, we examine more than 20,000 German radio sermons from 2012–2024. They range from 90-second sermons from *Kirche in 1Live* to 2.5-minute sermons on *HR4* and *SWR4* and to 4-minute sermons on *WDR4* (scan the QR code below for a sound sample). To get also an impression of the historical development, we have a corpus of 272 historical radio sermons (2300 pages) from the National Socialist era and 96 radio sermons (800 pages) from the GDR era.

4. Corpus Preparation: The digitization, machine-readable text preparation and enrichment with metadata and annotations of the radio sermons had to be solved differently for each time period. The solutions with their individual steps can be seen in fig. 1, examples in fig. 2–3.

5. Grammatical annotations of the radio sermons: The current corpus is enriched with annotations. To be able to analyze the sermons linguistically, each of the 10 million words must be tagged with its word class (POS tagging), each word form must be lemmatized, and syntactic structures such as nominal phrases or verbal phrases must be identified. We performed these annotations automatically in *CorpusExplorer*.

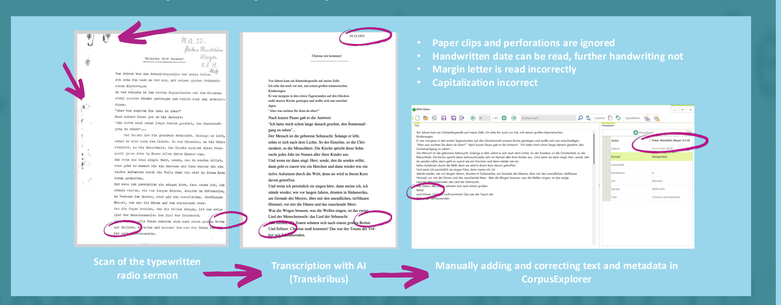

Fig. 2. **Example: Digitizing a historical radio sermon and making it machine-readable.**

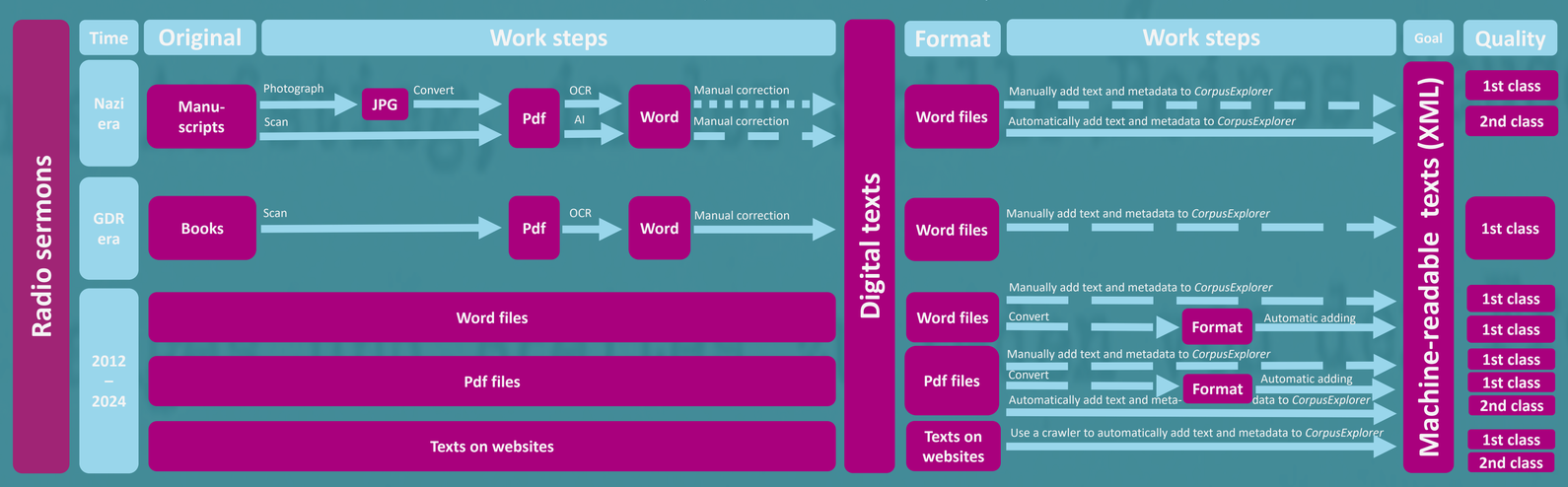

Fig. 1: **Visualization of the corpus preparation with its individual steps from the originals to the machine-readable texts with metadata.**

6. Analysis Methods: A mixed-methods approach is used that is both corpus-driven and corpus-based, combining quantitative and qualitative methods. First, the corpus is examined using quantitative analyses, without narrowly defined questions and concrete hypotheses. With machine support, frequently used words and patterns of linguistic acting can be identified. These are then used as a starting point for more in-depth quantitative and qualitative analyses, also using corpus-based methods. The radio sermons are thus analyzed corpus-pragmatically and qualitatively-hermeneutically. We use *CorpusExplorer*, a software that offers a wide range of methods for both types of analysis. For instance, we perform frequency, cooccurrence, keyword and N-gram analysis (see fig. 4 and 5).

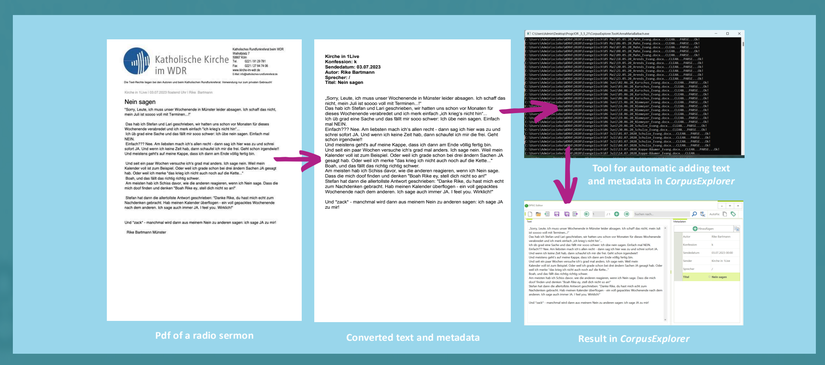

Fig. 3: **Example: Digitizing a modern radio sermon.**

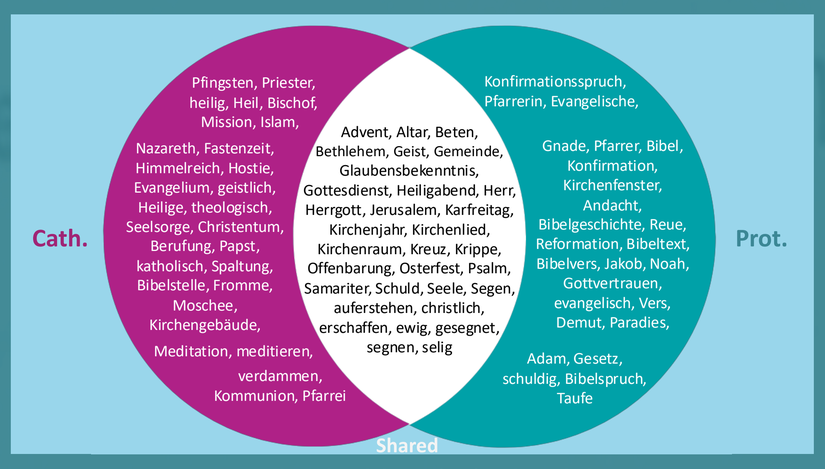

Fig. 4: **Shared and denominational vocabulary.**

7. Results:

Similarities between Catholic and Protestant Radio Sermons

- Large shared vocabulary
- Thematic focus: Talking about God, people, the relevance of God in their everyday life

Preferences of

Protestant Radio Sermons	Catholic Radio Sermons
• Bible stories, Bible quotations, people from the Bible	• Church, church holidays and feasts, veneration of saints and Mary, current events
• Frequency analyses: show frequent use of the noun "bible"	• Frequency analyses: show frequent use of the noun "church"
• Cooccurrence analyses: much biblical vocabulary, e.g., around "God"	• Cooccurrence analyses: "my church", church holidays and structures, critical statements
• Name analyses: Use of numerous personal names from the Bible	• Name analyses: Mary, saints' names, popes
• N-gram analyses: reveal language patterns to introduce biblical quotations or are biblical quotations	• N-gram analyses: Language patterns in 1st person

	3-grams		5-grams		7-grams	
	Prot.	Cath.	Prot.	Cath.	Prot.	Cath.
1	, dass ich	, dass ich	heißt es in der Bibel	Heute ist wieder Montag,	, wir feiern dich dafür, dass	Heute ist wieder Montag, und weil
2	, wenn ich	, wenn ich	", sage ich	ich weiß nicht, ob	vergiss nicht, was er dir Gutes	Es gibt ja immer was zu feiern
3	in der Bibel	, dass es	der Bibel heißt es:	seit Donnerstag im Kino läuft	reden über ihre Ängste und die	, der seit Donnerstag im Kino läuft
4	, dass es	Menschen, die	In der Bibel heißt es	, denke ich mir.	"Lobe den Herrn, meine Seele	"Liebe deinen Nächsten wie dich selbst
5	Menschen, die	, dass sie	", sagt sie.	über Gott und die Welt	Menschen, wenn er die ganze Welt	, dass es Menschen gibt, die
6	, dass sie	, dass die	Ich stelle mir vor,	ja immer was zu feiern	hat seinen Engeln befohlen, dass sie	selbst zu lieben begann , habe
7	Guten Morgen!	, was ich	", sagt er.	Menschen, mit denen ich	hat er wieder bei mir gekommen.	ich bin bei dir, wenn du
8	, was ich	, dass er	, heißt es in der	ich finde, das ist	"Liebe deinen Nächsten wie dich selbst	wenn ich das Gefühl habe, dass
9	, dass er	das, was	steht in der Bibel.	ich weiß, dass ich	Ihr Pfarrer Oliver Mahn aus Köln.	"Unser tägliches Brot gib uns heute
10	das, was	, dass wir	einer Zeit, in der	im wahrsten Sinne des Wortes	Don't fall, don't give up.	Liebt eure Feinde und betet für die
11	, dass die	, die ich	den Herrn , meine Seele	, habe ich mir gedacht	, wenn er die ganze Welt gewinnt	Gott, an den ich glaube,
12	, wenn man	, die sich	In der Bibel gibt es	ist noch nichts verloren.	Was glaubst du, wie oft	ich bin glücklich, auch wenn ich
13	", sagt	, in der	deinen Nächsten wie dich selbst	ich bin mir sicher,	bedenken, dass wir sterben müssen,	Was willst du, dass ich dir
14	, dass wir	, das ist	es in der Bibel:	deinen Nächsten wie dich selbst	nützt es einem Menschen, wenn er	", habe ich mir gedacht.
15	, in der	, dass das	In der Bibel steht:	, wie ich bin.	ist nicht gut, dass der Mensch	", hat er mir gesagt.

Fig. 5: **Examples of Protestant and Catholic N-grams.**

8. Conclusion: With the help of contrastive corpus linguistic, machine-based analyses and qualitative hermeneutic studies, we can demonstrate that even today, 500 years after the Reformation, denominational language differences still exist in certain fields. These are closely linked to the different theological orientations of the denominations and thus reflect linguistically different Catholic and Protestant orthodoxies.

9. Literature and our publications: Scan the QR-Code below.

1923 — ZENSIERT — 1924 — ZENSIERT — 1933 — 1945 — ZENSIERT — 1945–1990 — 2012–2022 — 2024

Kontakt:
Dr. Anna-Maria Balbach
Universität Münster, Germanistisches Institut
Schlossplatz 34, 48143 Münster
E-Mail: Anna.Balbach@uni-muenster.de

Literature: SCAN ME
Audio sample: SCAN ME
Instagram: SCAN ME — FOLLOW US

Funded by DFG Deutsche Forschungsgemeinschaft