

“What is Love ...?": A Study of Thematic Constructions in 11.955 Danish Sermons

Anne Agersnap/Kristoffer Laigaard Nielbo

Aarhus University



Anne Agersnap: <https://orcid.org/0000-0002-7135-7125>
Kristoffer Laigaard Nielbo: <https://orcid.org/0000-0002-5116-5070>

DOI: <https://doi.org/10.21827/ijh.5.1.86-116>

Abstract

Writing sermons in the Evangelical-Lutheran Church in Denmark (ELCD) is a shared practice, which means that pastors collectively produce a comprehensive text material, when they prepare their weekly sermons. This article studies a corpus of 11,955 ELCD sermons as a collective text production, as we investigate the role of theological discourses in emergent thematic fields in the corpus. With the aid of computational tools from the field of digital humanities, we investigate the following question: How can we map and interpret theological discourses in the collective production of sermons from the Evangelical-Lutheran Church of Denmark (ELCD)? Based on three complementary case studies, we explore the overall thematic framework of the corpus, the construction of specific theological concepts (love and sin) and the relationship between thematic constructions in the corpus and thematic content in the liturgical texts that pastors expound in sermons.

Keywords: sermons; collective practice; Danish Evangelical Lutheran Church of Denmark; large-scale text analysis; computational humanities

1. Introduction

When they preach on church holidays, pastors perform a significant interpretative task of making Christian narratives and teachings pertinent for assembled congregations in the present. As a consequence of this practice, pastors synchronously produce a large number of sermons on a weekly basis. In a Danish context, the vast majority of pastors write a full sermon manuscript, and these documents provide a unique insight into how pastors collectively on behalf of their church represent symbolic language and theological discourse. This knowledge potential has so far not been studied in sermons as outcomes of a collective practice. In this article, we therefore ask: *How can we map and interpret theological discourses in the collective*

production of sermons from the Evangelical-Lutheran Church of Denmark (ELCD)? We explore this question through three complementary case studies informed by the digital humanities, where we attend to thematic constructions in a collection of 11,955 ELCD sermons primarily from 2011–2016. In the first study, we calculate word frequencies to infer a general thematic framework of the corpus. In the second study, we map the semantics of the terms “love” and “sin” to explore how core theological concepts are constructed in a contemporary collection of sermons. In the third case study, we group the sermons in terms of holiday tides to investigate how thematic content from the liturgical year of the church aligns with thematic constructions in sermons. Through this bottom-up approach to digital text studies, the three cases in combination unveil different theological dynamics, which we discuss with a particular focus on the construction of “love” in the different cases.

Our interest in sermons stems from the sociological study of religion, but we consider our study to be in line with the empirical turn in homiletics, which has paved the way for descriptive analyses of preaching. So far, this turn has in particular been interested in the sermon as a component in a communicative situation, whether the focus has been on the speaker or more recently the listener¹. In our approach, we do not attend to the communicative intentions or consequences of preaching. Instead, we perceive the production of sermon documents as collective testimonies to Christianity and time.

2. The Evangelical-Lutheran Church in Denmark

The Evangelical-Lutheran Church in Denmark (ELCD) constitutes the majority religion in Denmark, and is affiliated with the state through section four in the Danish constitution, declaring: “The Evangelical-Lutheran Church is the Danish people’s church and as such supported by the state”.² The church is organised with ten dioceses divided geographically and supervised by ten bishops. Each diocese is divided into deaneries, and each deanery consists of local parishes. Membership in the ELCD is high, but in steady decline primarily due to demographic changes. On 1 January 2020, 74.3% of the Danish population were members compared to 80.4% on 1 January 2010. Despite high membership levels, regular attendance at church services is comparably low – about 11% of the population attend ELCD services at least

¹ Marlene Ringgaard Lorensen, Chapter 3: Dialogical Preaching: From ‘New Homiletics’ to ‘Other-wise Homiletics’, in: Dialogical Preaching: Bakhtin, Otherness and Homiletics, Göttingen 2013, 68–94.

² Translation from Danish: “§4. Den evangelisk-lutherske kirke er den danske folkekirke og understøttes som sådan af staten.”

once a month³. Meanwhile, the church has higher attendance rates in connection with ceremonies framing life events: in 2018, 58.7% of all newborns were baptised; in 2019, 68.3% of all youth in the valid age range had their confirmation; and of all the people who died in 2019, 82.7% had a church funeral⁴. The theology of the church is generally liberal: women pastors have been ordained since 1948, three out of ten bishops are women, and the performance of same-sex weddings has been allowed since 2012.

The current liturgy used at regular church services in the ELCD was approved in 1992 by the bishops at the time and authorised by the Danish Queen. Pastors are expected to follow this liturgy, but they can choose from a number of variations. The service includes two or three bible readings: from the Old Testament and/or Epistles and from the Gospels. The readings for each holiday are prescribed in two different lectionaries, one for odd and one for even liturgical years. Combined, these texts provide a thematic basis for each holiday. Although the two lectionaries are not identical, the order of texts represents the narrative of salvation throughout the course of a liturgical year⁵.

3. Sermons: Collective Testimonies to Christianity and Time

The empirical turn within homiletics has paved the way for analysing preaching inductively. This development has largely entailed a shift from a kerygmatic to a communicative understanding of preaching⁶. Instead of regarding preaching firstly as an event of divine proclamation uttered by the preacher, the empirical turn has increasingly focused on the listeners' experience of sermons. Danish theologians Marlene Ringgaard Lorensen and Kirstine Helboe Johansen have outlined the status of empirical research within the field of homiletics in the Nordic countries⁷. Here, the strongest tendencies draw inspiration from the newer North American tradition known as *Other-wise Preaching*, which has been developed in particular in the works of John McClure⁸. In *Other-wise preaching*, focus is on the

³ Peter B. Andersen, Jakob Erkman & Peter Gundelach, Udviklingen i (ikke)religiøsitet, in: Usikker modernitet: Danskernes værdier fra 1981 til 2017, ed. Morten Frederiksen, Copenhagen 2019, 231–264.

⁴ All statistics can be found at: <https://www.km.dk/folkekirken/kirkestatistik/>

⁵ Henning Thomsen, Prædikenens plads og indhold, in: Pastoralteologi, eds. Helle Christiansen and Henning Thomsen, Copenhagen 2007, 260–269.

⁶ F. Gerrit Immink, Homiletics: The Current Debate, in: International Journal of Practical Theology 8.1 (2004), 89–121.

⁷ Marlene Ringgaard Lorensen & Kirstine Helboe Johansen, Homiletics in the Nordic Countries, in: Praktische Theologie 55.2 (2020), 88–93.

⁸ John McClure, The roundtable pulpit: where leadership and preaching meet, Nashville 1995; John McClure, Other-wise preaching: a postmodern ethic for homiletics, St. Louis 2001.

asymmetrical relationship between preacher and listener, meaning that the preacher cannot assume what congregants learn from the sermon, as the congregants come with backgrounds and perspectives different from those of the preacher. In a Nordic context, this approach to sermons has informed studies with an increasing focus on what congregants hear and interpret from the sermon. In Denmark, homiletician and now bishop Marianne Gaarden has studied this phenomenon in her doctoral thesis, arguing that listening to a sermon creates a third room between speaker and listener⁹. Similarly, in Norway, a collective of researchers has studied how listeners at different ages experience sermons¹⁰, and in Sweden, homiletician Jonny Karlsson has examined how different congregations have been addressed in the sermon collection of theologian Gustaf Wingren¹¹. With such studies, the homiletical discipline has become subject to empirical research with an interest in not only what preaching ought to be, but also how preaching practices actually unfold.

The empirical turn in the Nordic countries has primarily entailed studies of sermons with regard to the communicative situation in church between preacher and congregant. Yet, sermons are not merely subjective realities – not just the intentional proclamation of the preacher and the experienced perception of the listener. Sermons take part in a collective voice, which is documented by the thousands of sermons pastors write every week¹². Literary scholar Franco Moretti uses the term *collective* to describe the relations between literary works of the same historical period. His description illustrates a shift from approaching texts as unique individual works to understanding them as parts of a larger literary field. Moretti argues: “a field this large cannot be understood by stitching together separate bits of knowledge about individual cases, because it *isn't* the sum of individual cases: it's a collective system, that should be grasped as such, as a whole”¹³. Therefore, a literary work is not just the product of an author's creative mind, but also a product conceived in a specific cultural and temporal context in history. As such, it gives access to a collective consciousness that may bypass the author's own intention or awareness – a collective consciousness that we can only study if we examine the individual work as intertextually related to the larger production of works sharing historical context. This perspective in itself is not unique to the field of literary

⁹ Marianne Gaarden, *The Third Room of Preaching: The Sermon, the Listener, and the Creation of Meaning*, Louisville 2017.

¹⁰ See the project web-page: <https://www.mf.no/kom/forskning/foss-forkynnelse-sma-store>

¹¹ Jonny Karlsson, *Predikans samtal: en studie av lyssnarens roll i predikan hos Gustaf Wingren utifrån Michail Bachtins teori om dialogicitet*, Linköping University 2000.

¹² Anne Agersnap, Ross Deans Kristensen-McLachlan et al., Sermons as data: Introducing a corpus of 11,955 Danish sermons, in: *Journal of Cultural Analytics* 12 (2020), 1–27.

¹³ Franco Moretti, *Graphs, Maps, Trees – Abstract Models for Literary History*, London/New York 2007, 3–4.

history, but Moretti radicalises the approach: even though literary historians may study hundreds of works of a given period, these collections tend to comprise canonised works, which means that the same “few” literary works are scrutinised closely over and over again by different scholars¹⁴. Instead of reading only a few works, Moretti’s concept of *distant reading* implies adopting a wider view on literature, by attending to a larger collection of works from a given period in time¹⁵. The idea of distant reading has become important for the field of digital humanities, which since then has contributed with a range of computational tools for large-scale text studies.

Moretti’s perspective on literature as embedded in a collective system provides new ways of embarking on empirical research of sermons. In parallel to literary works, sermons are conceived in a particular temporal and cultural context, where certain discourses and ideas prevail, and their shared contemporaneity situates the sermons in an intertextual relationship. Furthermore, even more so than for literary works, sermons are embedded in a concrete system due to the liturgical years, whereby sermons are produced in parallel to each other and on a continuous, recurring basis. In a Danish context, the biblical passages prescribed for each holiday in fact provide a tangible intertextual basis for sermons. When we transpose Moretti’s idea of collective text systems to the context of sermons, we therefore define the collective production of sermons as a text system that is intricately embedded in Christian tradition as well as contemporary culture. Furthermore, we argue that traits of these religious and cultural codes can be observed through large-scale text studies¹⁶.

The interest in studying sermons in the intersection between theological and contemporary public discourses is not unique in a Nordic context. For example, a research collective including Lorensen and Norwegian homiletician Tone Stangeland Kaufman has examined the influence of the international refugee crisis in 2015 on preaching in Denmark, Norway and Sweden¹⁷. The study was done on 104 sermons delivered over the course of six months – an impressive amount of text for a close-reading study. However, with Moretti’s more radical understanding of collective text productions, we can broaden the scope of study even more through the aid of a large digital archive of sermons and computational methods. By approaching sermons as a large collective production of texts by pastors, we leave out the

¹⁴ Ibid., 3–4.

¹⁵ Idem, Conjectures on world literature, in: *New Left Review* 1 (2000), 54–68.

¹⁶ The understanding of sermons as a collective text system is developed further in: *Anne Agersnap, Collective Testimonies to Christianity and Time*, PhD dissertation, Aarhus University 2021, 20–37.

¹⁷ *Marlene Ringgaard Lorensen, Tone Stangeland Kaufman et al., Preaching in times of the European ‘Refugee Crisis’. Scandinavian perspectives*, in: *International Journal of Homiletics* 1.2 (2017), 74–100.

communicative purpose and impact of individual orally delivered sermons¹⁸. What is to be gained, however, is an insight into sermons as collective testimonies to Christianity and time. As such, we are able to investigate prevalent themes, and how theological vocabulary and discourse unfold in a temporally coherent collection of sermons.

4. Transformational Patterns in Contemporary Christianity

Our approach to conducting semantic investigations of thematic content and theological dynamics in sermons is inspired by sociologist Linda Woodhead’s concept of the *displacement of theology*. She uses the concept to explain how theological expertise has been devalued in public life in European societies, in particular since the 1960s. Due to new power structures, theology as an academic discipline lost both social power and cultural influence. In society, the influence of Christianity thus seems to have shifted from a ‘power from above’ (confessional theology and authoritative dogmas) to ‘a power from below’ (lay peoples’ demands to religion)¹⁹. Recent studies have investigated what these developments mean for the practices in the ELCD. Marie Vejrur Nielsen and Kirstine Helboe Johansen demonstrate how new types of church services, such as Halloween and Valentine’s Day services, are situated in a church context while clearly drawing on mainstream cultural codes in framing them²⁰. Another study conducted by Nielsen focusing on the profile and motivations of the users of the ELCD shows that a majority of the informants are occasional users, who are interested in attending specific activities in their church, for example baby hymn singing for parents and babies, rather than regular attendance at church services²¹. These studies indicate that theological displacements occur in the ELCD, as churches provide new activities that tend to meet the demands, needs or interests of its members rather than providing a setting for intellectual theological discussion. Since these activities are supplementary to the official church calendar and not bound to the liturgical order of church services, they allow a higher degree of diversion from intellectual or doctrinal theological discourse in order to reach specific target groups. For the weekly and

¹⁸ Moreover, we are aware that pastors may not have delivered the sermons in church as written in the sermon manuscripts. The document content, however, is still an attestation of pastors’ theological work today, whether pastors’ oral presentation of the sermon reflects the content as stated in the document or not. The sermon document is a concrete imprint that has been made before the service, and a concrete imprint that remains after it.

¹⁹ Linda Woodhead, *An Introduction to Christianity*, Cambridge 2004, 333–347, 361–366.

²⁰ Marie Vejrur Nielsen & Kirstine Helboe Johansen, *Transforming churches: the lived religion of religious organizations in a contemporary context*, in: *Journal of Contemporary Religion* 34 (2019), 509–527.

²¹ Marie Vejrur Nielsen, *Changing patterns? Occasional consumers of new activities in old churches*, in: *Nordic Journal of Religion and Society* 28.2 (2015), 137–153.

authorised church services, the situation is different. The sermon for the official church holidays is a genre that has traditionally been a distinct domain of theological expertise. Where Woodhead uses the term theological displacement to describe a societal dynamic, and where Nielsen and Johansen look at dynamics on an organisational level by attending to supplementary church activities, we investigate which theological dynamics occur in one of the church’s core practices shared by all pastors in the ELCD. When we define ELCD sermons as a collective text system intricately embedded in theological and cultural discourses, the idea of theological displacements become interesting for exploring theological dynamics in this text production. A strict use of Woodhead’s concept would entail, that cultural discourses trump theological ones. In our approach, however, we remain open towards the idea that cultural codes and theological codes may interact as well as preclude one and other. To explore this, we look for thematic content inductively in a large sermon corpus. This approach entails paying close attention to pastors’ shared vocabulary, the configurations of the vocabulary and the role of theological language in these configurations. Through these patterns, we are able to explore the collective voice of ELCD pastors.

5. Methods and Observations

Data and Computational Tools

The corpus we analyse consists of 11,955 sermons from the ELCD. Each sermon is annotated with metadata, including pastor pseudonym, date, holiday, parish size, diocese, pastors’ gender, pastors’ year of birth and pastors’ place of education. The sermons were submitted by 95 pastors, and they are predominantly from the period 2011–2016 (roughly 96% of the corpus), while a few sermons are from before or after this period or are lacking information about the date. The sermons are collected based on opportunity sampling, not probability sampling. The corpus does therefore not constitute a representative sample of sermons in the ELCD, but instead, it forms a comprehensive case²².

In our approach to studying themes in sermons, we attend to word representations in the text collection and explore the semantic contexts they produce and are part of. In this endeavour, we model lexical semantics in terms of the co-occurrence and associative structure of the corpus at the level of individual sentences, documents and the corpus as a whole. For the analysis, we are less interested in the concrete syntactic context of words in individual

²² For a thorough presentation of the sermon corpus (sampling, metadata and basic corpus statistics), see: *Anne Agersnap, Ross Deans Kristensen-McLachlan et al., Sermons as data: Introducing a corpus of 11,955 Danish sermons*, in: *Journal of Cultural Analytics* 12 (2020), 1–27.

sermons; instead we focus on semantics in terms of lexical centrality, repetition and relations across the entire corpus.

We have applied three tools for our semantic investigations: *word count analysis*, *topic modelling* and *seeded semantic network analysis*. Our design is thus complementary with three case studies that combined provide an exploratory approach to investigate and trace thematic content in the corpus²³. Word count analyses provide an overview of the vocabulary of the corpus in terms of word frequencies. We thus consider a high frequency of some words rather than others as an indicator of central status in the collection. As we attend to the high frequency words in relation to each other, we get an immediate sense of some of the important themes in the collection. Topic modelling is a tool to automatically represent latent themes that run through a large collection of texts. A topic model makes the assumption that a document is a distribution over a finite set of latent topics, and that a latent topic is a distribution over the corpus lexicon. Furthermore, it makes the assumption that the topics within a text have been chosen by the author before writing the text, and that the words within a text are realisations of these pre-existing topics. With topic modelling, you try to reverse this process and trace the words back to the topics they came from, and consider the documents in terms of their distribution of topics. This can be done using a probabilistic approach, where an algorithm computes the probability that every word within every document belongs to any of the number of topics you are looking for. Thus, in topic modelling, you define how many topics it is reasonable to look for within a given text collection before conducting the analysis²⁴. Topic modelling regards words to be central to a given theme by considering how each word in the text collection occurs within every document of the collection. This approach implies relating documents to each other based on their distribution of topics. In contrast to word count analyses and topic modelling, a seeded semantic network queries the associative structure of concrete seed terms in the corpus. This technique uses simple neural word embeddings to represent the corpus lexicon and builds an associative graph by using words as nodes and distance in word context as edges²⁵. By iteratively applying the algorithm with threshold for graph size, the technique builds a hierarchical graph, where the primary associative level

²³ The three case studies in combination allow for an inductive investigation of semantic fields, even though some of the pipelines in the individual cases include deductive phases.

²⁴ David M. Blei, Probabilistic Topic Models, in: Communications of the ACM 55.4 (2012), 77–84.

²⁵ T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean, Distributed representations of words and phrases and their compositionality, in: Advances in Neural Information Processing Systems (2013) 3111–3119; C. Martinez-Ortiz, T. Kenter, M. Wevers, P. Huijnen, and J. van Eijnatten, Design and implementation of ShiCo: Visualising shifting concepts over time, in: Proceedings of the 3rd HistoInformatics I Workshop (2016), eds. M. Düring, A. Jatowt, J. Preiser-Kapeller, A. van den Bosch, 9.

consists of the m most similar terms to the seed term, and the secondary associative level consists of the n most similar terms to the primary associative level. Finally, the edges for all terms of the graph are computed with a distance threshold, and the Louvain method for community detection is used to extract semantic clusters²⁶. In our terminology, the repetitive and relational structures that these tools uncover establish semantic fields that can be scrutinised closely. The words in the sermon collection are thus understood as text components facilitating links between the texts. These links exhibit latent patterns that would not be possible to uncover with the naked eye and without the aid of computational tools. This way, our approach allows us to explore thematic content in the full corpus and thereby to study it as a collective and intertextual text system.

Pipeline: Word Count

For our word count analysis, we tokenised and casefolded the entire corpus. We then calculated how frequently identical tokens, i.e. words, appear in the text collection. There are in total 14,381,998 tokens in the corpus, and 125,311 unique tokens, which we call terms. This calculation provides a list of the most frequently mentioned words. The list shows that the most frequent words (MFW) in particular are stopwords, which are generic or function words – such as “and”, “that”, “to”, “we”, “there”. This type of words are innate terms for constructing most written texts despite genre. We filtered out such stopwords, leaving only nouns and proper nouns on the list, in order to represent terms that carry semantically significant content in this text collection. These terms are presented in Table 1 below.

²⁶ Vincent D. Blondel et al., Fast unfolding of communities in large networks, in: *Journal of Statistical Mechanics: Theory and Experiment* vol. 2008 no. 10 (2008), P10008, doi: [10.1088/1742-5468/2008/10/P10008](https://doi.org/10.1088/1742-5468/2008/10/P10008).

Term	Frequency rank	Frequency score
god (<i>gud</i>)	26	96,278
jesus (<i>jesus</i>)	27	88,555
god's (<i>guds</i>)	45	45,269
life (<i>liv</i>)	48	43,227
day (<i>dag</i>)	56	32,523
human beings (<i>mennesker</i>)	57	32,307
the life (<i>livet</i>)	72	26,047
world (<i>verden</i>)	74	25,732
word/words (<i>ord</i>)	78	24,699
faith (<i>tro</i>)	82	23,906
love (<i>kærlighed</i>)	98	19,769
human being (<i>menneske</i>)	108	17,273
time (<i>tid</i>)	119	15,511
jesu (<i>jesu</i>)	120	15,507
christ (<i>kristus</i>)	132	13,401
year (<i>år</i>)	142	12,401
son (<i>søn</i>)	151	11,364
children (<i>børn</i>)	155	10,916
joy (<i>glæde</i>)	163	10,428
light (<i>lys</i>)	170	9,659
thing (<i>ting</i>)	172	9,491
realm/rich (<i>rige</i>)	173	9,358
path (<i>vej</i>)	174	9,310
place (<i>sted</i>)	176	9,138
death/dead (<i>død</i>)	177	9,112
child (<i>barn</i>)	178	9,081
man (<i>mand</i>)	180	8,955
the faith (<i>troen</i>)	186	8,448
people (<i>folk</i>)	188	8,383
hope (<i>håb</i>)	189	8,349
earth (<i>jorden</i>)	190	8,283
the church (<i>kirken</i>)	191	8,245
the death (<i>døden</i>)	192	8,221

Table 1: List of the most frequent content words in the sermon corpus. The table describes the terms' frequency rank and frequency scores. Terms with ambiguous translations are reported with a dash (/). Terms in parenthesis are the Danish term.

Pipeline: Seeded Semantic Network Analysis

We use a seeded semantic network to trace the associative structure of the terms “love” and “sin”. “Love” is high on the MFW list (the 98th most frequent word) indicating that it is an important concept in the sermon corpus. In the context of contemporary sermons, this concept is furthermore interesting as it on the one hand has strong theological connotations, while it on the other hand is a common and important concept in contemporary culture. With the seeded semantic analysis, we are able to explore the more complex connotations of the term in the corpus and its relation to theological and cultural codes. To contrast the construction of “love”, we investigate the semantics of “sin”, which is also an inherent part of the Christian symbolic system, but comparably less disseminated in the corpus (the 492nd most frequent word) and without the same mainstream and positive significance in contemporary culture. With the semantic network analysis, we analyse and visually inspect the semantic constructions that these two terms tend to be part of, and how the terms relate to each other. We thereby have a tool to query targeted terms in the corpus with the purpose of examining conceptualisations bottom-up.

To build a seeded semantic graph, we initially train simple neural word embeddings using the CBOW model architecture²⁷ to learn distributed lexical representations. To compensate for the relatively small corpus size, the model is initialised on pre-trained weights for Danish. In order to build the graph, the algorithm first computes the distance between a set of seed terms, “love” and “sin”, and the corpus lexicon using the inverse trigonometric arccosine function. For each seed, the algorithm then extracts the m terms with the shortest distance, the primary associations, and then the n terms with the shortest distance, the secondary associations, to the primary associations. The distance between all terms (i.e., seeds, primary and secondary associations) are computed and terms are connected based on their distance under a given threshold (thresholds are estimated from the distance variance structure). Finally, semantic clusters are extracted using the Louvain method, and the graph is represented visually with terms as nodes and thresholded distances as edges. For visualisation purposes, seed and primary associations are represented with UPPERCASE and node colour indicates cluster²⁸.

²⁷ T. Mikolov, K. Chen, G. Corrado and J. Dean, Efficient Estimation of Word Representations in Vector Space, in: arXiv:1301.3781 [cs], Sep. 2013. Accessed: Jan. 05, 2021. [Online]. Available: [1301.3781] [Efficient Estimation of Word Representations in Vector Space \(arxiv.org\)](https://arxiv.org/abs/1301.3781).

²⁸ Code applied for the seeded semantic network analysis is available here: [GitHub - centre-for-humanities-computing/Semantic-Kernel: Tool for building and visualizing neural concept graphs](https://github.com/centre-for-humanities-computing/Semantic-Kernel).

We represent the associative structures of “love” and “sin” in three graphs. In Figure 1 and Figure 2 below, we have the semantic network of “love” and “sin” individually²⁹.

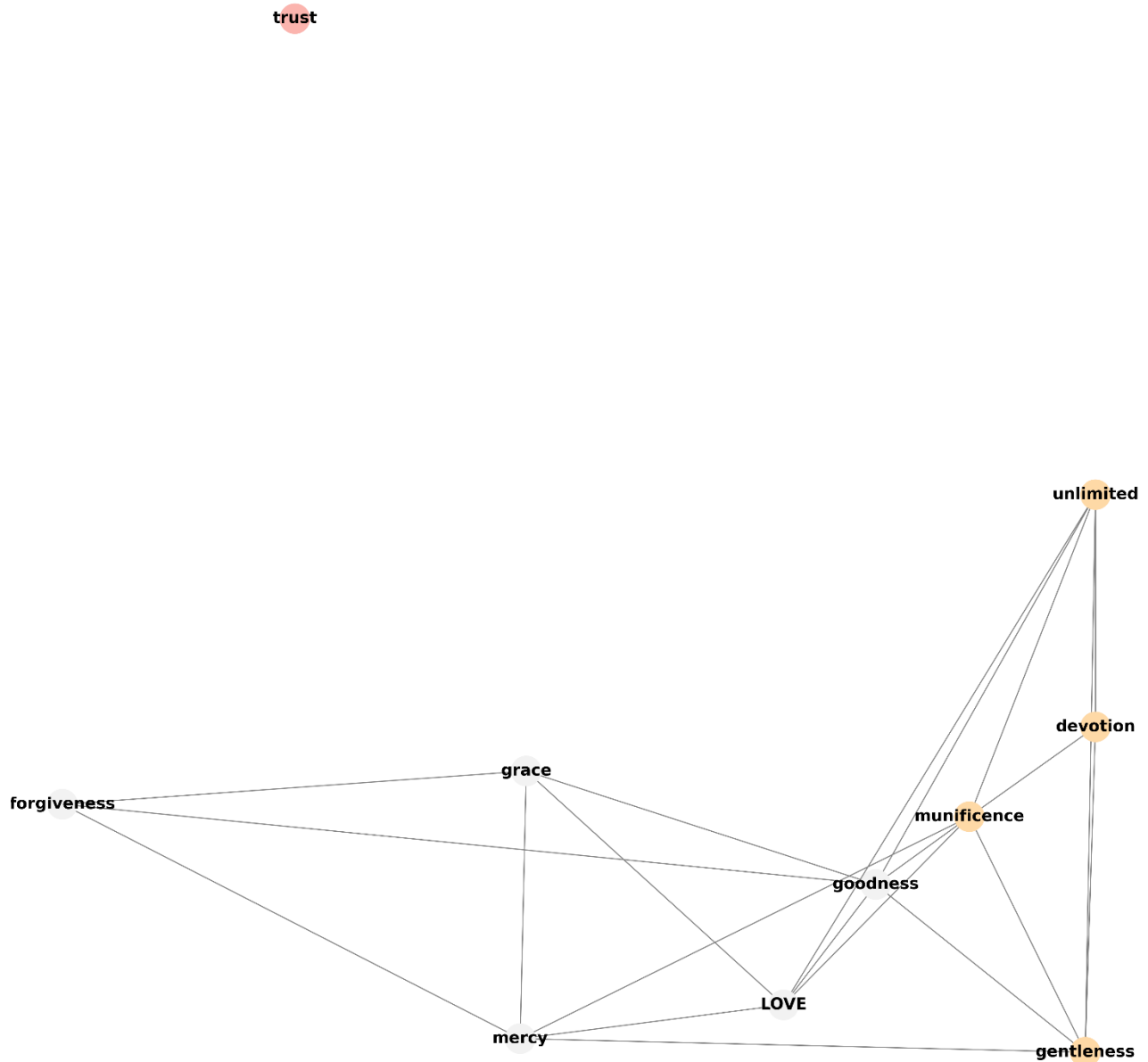


Figure 1. Semantic network of the seed term “love” and primary associations.

²⁹ All terms in the semantic networks are the authors’ English translations of the original observations. See the original Danish terms in Appendix A.

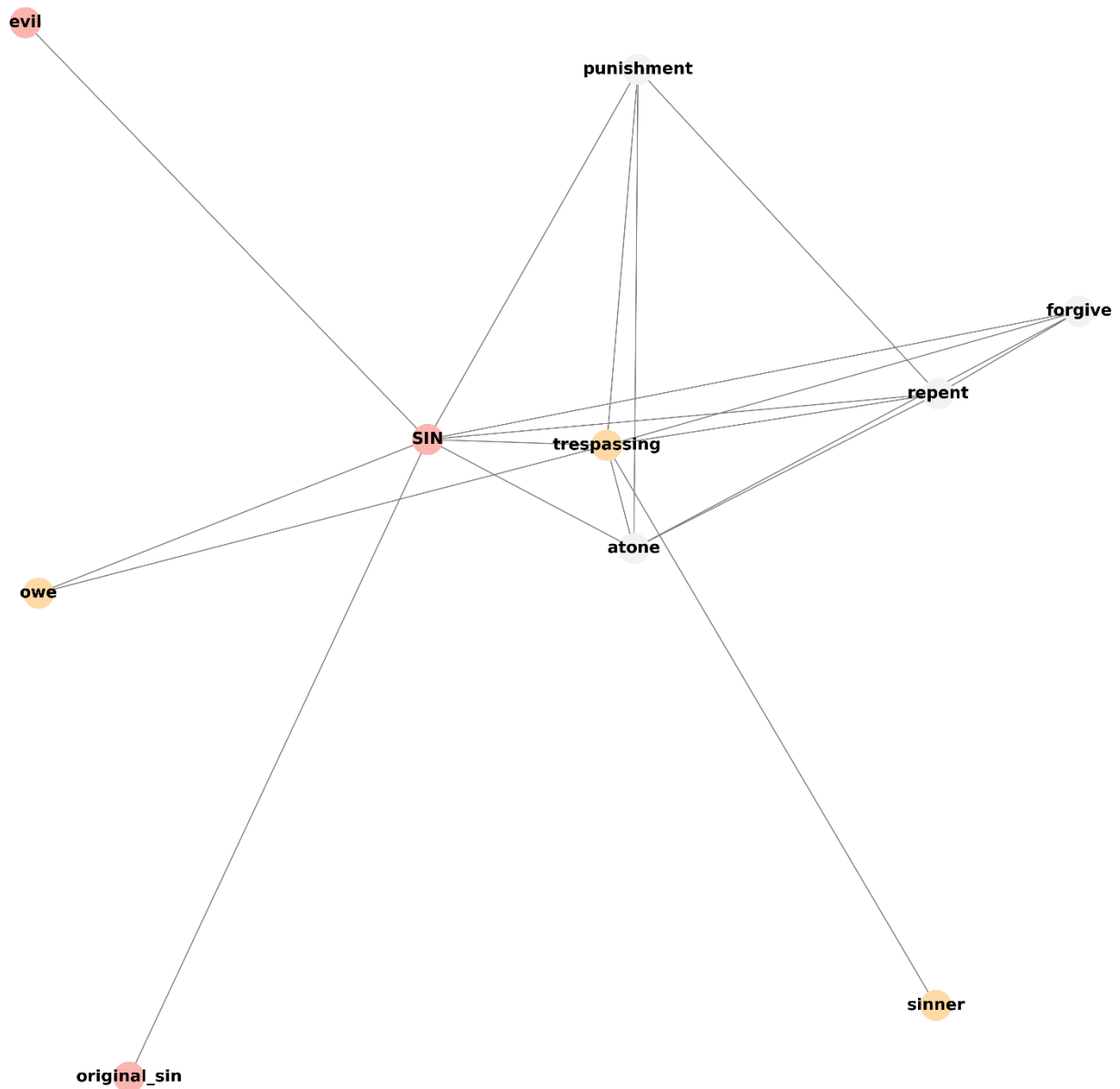


Figure 2. Semantic network of the seed term “sin” and primary associations.

In Figure 3 below, we represent the semantic networks in relation to each other, where the grey nodes represent terms associated with “love”, and the pink nodes are terms associated with “sin”. Each seed in Figure 3 appears with ten primary associations, and each primary association has three secondary associations.

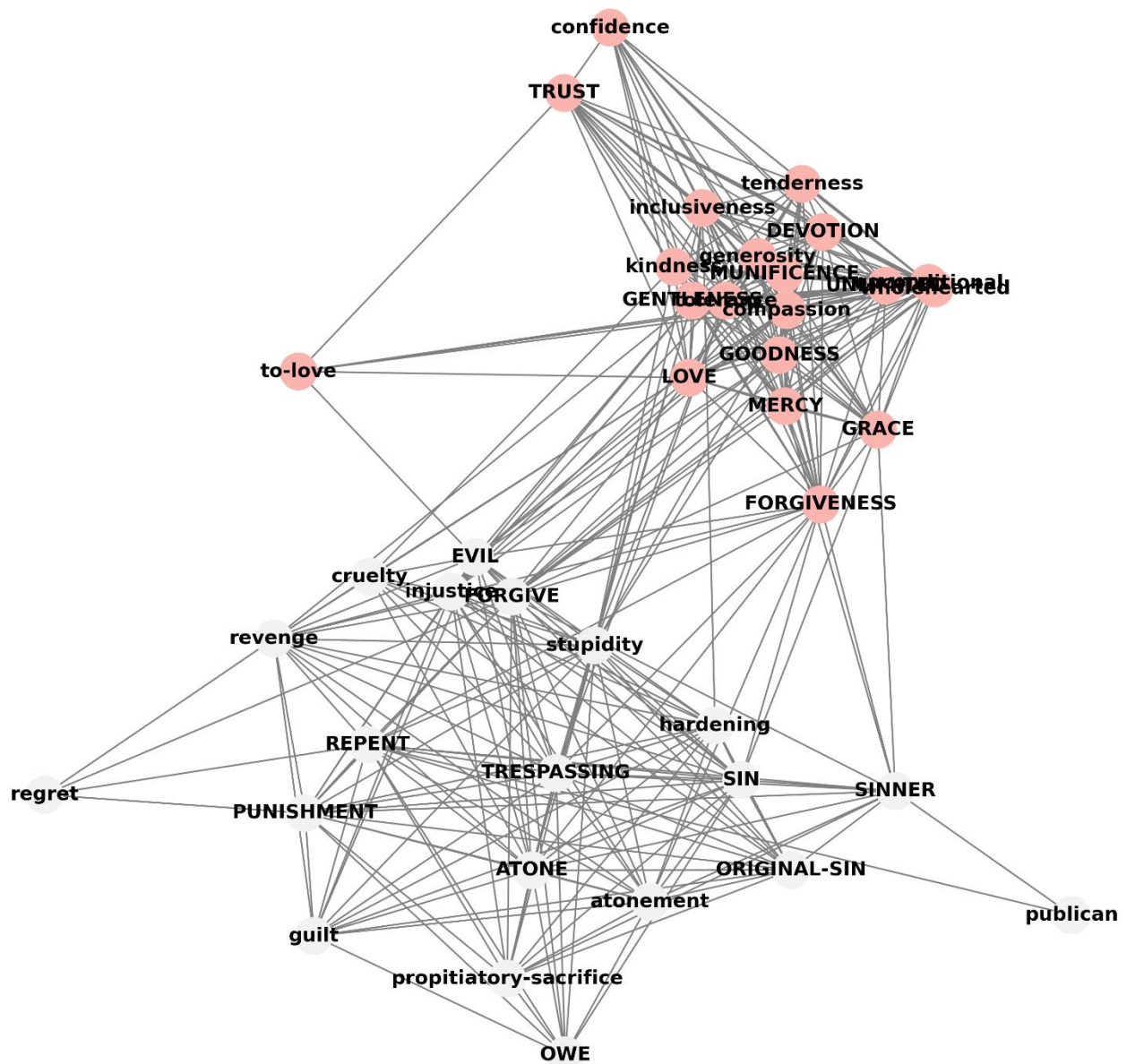


Figure 3: Semantic network of the seed terms “love” and “sin” and primary and secondary associations

Pipeline: Topic Modelling

Topic modelling is a bottom-up approach to classify thematic content. However, we designed our topic modelling analysis as a hypothesis-deductive test of whether holidays in the liturgical year explain thematic content in sermons, asking: do sermon documents from the same holidays share the same themes? For each of the 95 pastors, we chose to concatenate their sermons according to one of ten holiday periods in the church year, and we did therefore not run the topic model on 11,955 individual documents³⁰. Before the concatenation, we divided the holidays of the church year into ten sub-groups, which we defined. Our criteria for this division were that holidays named according to same period in the liturgical year should belong to the same sub-group; that the holidays in the same sub-groups should follow each other consecutively according to the liturgical year; and that the sub-groups should be as even in size as possible, while complying with the first two criteria. See the holiday sub-groups in Appendix B.

As the time of Trinity extends over five months in the liturgical year, we divided this season into three different sub-groups. Similarly, the holidays after Easter are divided between two groups, “POST-EAST_1” and “POST-EAST_2”, even though these groups contain holidays that actually represent a delimited holiday period themselves (for example Pentecost). However, due to our criteria, they do not constitute their own group in our division. Based on the holiday metadata in our corpus, we concatenated the pastors’ sermon productions according to the ten holiday periods in Appendix B. Our new corpus thus holds up to ten concatenated documents from each pastor, but in practice, we ended up with a corpus of 871 documents. Therefore, some pastors are represented with fewer documents than ten. The documents are named according to holiday sub-group and the random pseudonym the pastors have in the metadata³¹.

We used DARIAH’s topics explorer, a user-friendly application for performing topic modelling, to analyse our new corpus³². The application tokenised and lowercased every word in the corpus and removed stopwords. We used a stopwords list that removed 262 function words from the corpus³³. We did not lemmatise the corpus, because we did not want to remove

³⁰ Our reasons for concatenating sermons were as follows: It would be too immense to survey this many documents in relation to all 69 holidays in the liturgical year; the documents in the sermon corpus differ in length, and some documents would therefore not provide enough data for the model to work adequately; the pastors had not provided sermons for every holiday in the liturgical year throughout 2011–2016, and we did therefore not have an even amount of data from each pastor.

³¹ For example, the document “13_Advent” contains the concatenated advent sermons from pastor number 13.

³² The application is available here: <https://dariah-de.github.io/TopicsExplorer/>

³³ The stopwords list is available here: <https://gist.github.com/berteltorp/ocf8aoc7afea7f25ed754f24cfc2467b>

possible nuances that different forms of a word can provide in the topics we extracted. We tested the application with 100 iterations, asking for, respectively, 10, 20, 50 and 100 topics. We report our results based on findings from 20 topics, as the tendencies from each test follow a similar pattern in terms of how documents are grouped based on the topics. Our analysis shows that the holiday signal of the sermons is most significant for how topics are distributed in the documents: among the 20 topics, 11 indicate a strong holiday signal, with eight or more documents adhering to the same holiday period. See Table 2 below for a distribution of topics in relation to holidays. Thus, each of our ten holiday periods is represented at least once in our topic modelling, while TRINITY_1 is represented in relation to two topics (topic 3 and 10). In comparison, we find five topics where the pastor signal is strongest with eight or more documents originating from the same pastors. Among the 20 topics, topics 2, 5, 7, 13 and 14 are most frequently represented in the sermons of individual pastors. Topics 0, 12, 15 and 19 reveal neither a strong holiday nor a strong pastor signal³⁴.

³⁴ All topics are enclosed in Appendix C.

Holiday signal	Topic	Top 15 words
ADVENT	Topic 17	john (<i>johannes</i>), jesus, day (<i>dag</i>), time (<i>tid</i>), light (<i>lys</i>), god’s (<i>guds</i>), advent (<i>advent</i>), christ (<i>kristus</i>), zechariah (<i>zakarias</i>), god (<i>gud</i>), joy (<i>glæde</i>), path (<i>vej</i>), awaits (<i>venter</i>), people (<i>folk</i>), isaiah (<i>esajas</i>)
CHRIST-NEW	Topic 16	god (<i>gud</i>), christmas (<i>julen</i>), child (<i>barn</i>), god’s (<i>guds</i>), world (<i>verden</i>), christmas (<i>jul</i>), year (<i>år</i>), light (<i>lys</i>), born (<i>født</i>), the child (<i>barnet</i>), peace (<i>fred</i>), human beings (<i>mennesker</i>), joy (<i>glæde</i>), bethlehem, joseph (<i>josef</i>)
EPIPHANY	Topic 9	small (<i>små</i>), old (<i>gamle</i>), men (<i>mænd</i>), children (<i>børn</i>), water (<i>vand</i>), place (<i>sted</i>), child (<i>barn</i>), wise/show (<i>vise</i>), become (<i>blevet</i>), adults (<i>voksne</i>), exactly (<i>netop</i>), at the time (<i>dengang</i>), narrative (<i>fortælling</i>), heaven (<i>himlen</i>), earth (<i>jorden</i>)
LENT	Topic 1	jesus, god (<i>gud</i>), god’s (<i>guds</i>), belief (<i>tro</i>), mary (<i>maria</i>), rich/realm (<i>rige</i>), day (<i>dag</i>), human beings (<i>mennesker</i>), son (<i>søn</i>), bread/broke (<i>brød</i>), jesus, human being (<i>menneske</i>), power (<i>magt</i>), evil (<i>onde</i>), believe (<i>tror</i>)
EASTER	Topic 6	jesus, dead/death (<i>død</i>), jesus, “the” death (<i>døden</i>), resurrection (<i>opstandelse</i>), life (<i>liv</i>), jerusalem, died/the dead (<i>døde</i>), the tomb (<i>graven</i>), the disciples (<i>disciplene</i>), christ (<i>kristus</i>), easter morning (<i>påskemorgen</i>), good friday (<i>langfredag</i>), the cross (<i>korset</i>), son (<i>søn</i>)
POST-EAST_1	Topic 4	jesus, peter, god (<i>gud</i>), shepherd (<i>hyrde</i>), life (<i>liv</i>), thomas, the truth (<i>sandheden</i>), good (<i>gode</i>), say (<i>siger</i>), freedom (<i>frihed</i>), day (<i>dag</i>), the path (<i>vejen</i>), the disciples (<i>disciplene</i>), recognise (<i>kender</i>), time (<i>tid</i>)
POST-EAST_2	Topic 8	god (<i>gud</i>), god’s (<i>guds</i>), jesus, word/words (<i>ord</i>), jesus, christ (<i>kristus</i>), the holy spirit (<i>helligånden</i>), life (<i>liv</i>), spirit (<i>ånd</i>), world (<i>verden</i>), human beings (<i>mennesker</i>), prayer (<i>bøn</i>), christian (<i>kristne</i>), community (<i>fællesskab</i>), christi (<i>kristi</i>)

TRINITY_1	Topic 3	jesus, god (<i>gud</i>), rich/realms (<i>rige</i>), god's (<i>guds</i>), life (<i>liv</i>), human beings (<i>mennesker</i>), peter, love (<i>kærlighed</i>), say (<i>siger</i>), son (<i>søn</i>), jesu, to love (<i>elske</i>), the parable (<i>lignelsen</i>), good (<i>gode</i>), money (<i>penge</i>)
TRINITY_1	Topic 10	say (<i>siger</i>), right (<i>ret</i>), received (<i>fået</i>), guilt (<i>skyld</i>), said (<i>sagde</i>), place (<i>plads</i>), go (<i>gå</i>), well (<i>vel</i>), to live (<i>leve</i>), day (<i>dag</i>), give (<i>give</i>), act (<i>handler</i>), whole (<i>hele</i>), children (<i>børn</i>), willingly (<i>gerne</i>)
TRINITY_2	Topic 11	jesus, god (<i>gud</i>), life (<i>liv</i>), word/words (<i>ord</i>), god's (<i>guds</i>), martha, day (<i>dag</i>), mary (<i>maria</i>), say (<i>siger</i>), thanks (<i>tak</i>), jesu, human beings (<i>mennesker</i>), lazarus, death (<i>døden</i>), good (<i>gode</i>)
TRINITY_3	Topic 18	god (<i>gud</i>), jesus, god's (<i>guds</i>), life (<i>liv</i>), the life (<i>livet</i>), human beings (<i>mennesker</i>), belief (<i>tro</i>), day (<i>dag</i>), give (<i>giver</i>), death (<i>døden</i>), to give (<i>give</i>), each other (<i>hinanden</i>), good (<i>gode</i>), jesu, human being (<i>menneske</i>)

Table 2: Relationship between holidays and topic. The table exhibits the holiday signals found from the generated topics and the 15 most related words within each topic. The original Danish terms are in parentheses, and the dash (/) indicates ambiguous meaning in the translation from the Danish terms. Topic numbers are assigned randomly to the topics by the topic model.

Our findings indicate that the structure of the liturgical years influences thematic content in sermons. We thus have a model for exploring how themes and concepts unfold in relation to the liturgical year, even though we used a hypothesis-deductive approach to infer Table 2.

6. Thematic Contexts in Sermons

The Large Framework: Solitude for Human Life

Our MFW list of content words, Table 1, indicates a rather coherent thematic framework for the sermon corpus. “God” and “Jesus” are the most frequent content words, which convincingly points out them as the main figures. The succeeding content words would emphasise a focus on a this-worldly life (“life”, “the world”, “earth”, “death”) among people (“human beings”, “children”, “man”, “people”) living here and now (“day”, “time”, “year”) from an optimistic perspective (“belief”, “love”, “joy”, “light”, “hope”). The conceptual framework found from the MFW list seems to illustrate a general theme about care for human life in the here-and-now, through a vocabulary that includes many terms that do not only have Christian connotations. It is striking that many concepts pointing to dogmas and cosmology are absent from the MFW list. The frequency rank (FR) of such terms is considerably lower – for example “heaven” (FR: 283), “the holy spirit” (FR: 378), “forgiveness” (FR: 387), “resurrection” (FR: 398), “mercy” (FR: 400), “redemption” (FR: 440), “the cross” (FR: 460), “sin” (FR: 492), “judgement” (FR: 672), “the devil” (FR: 742), “evil” (FR: 888), “hell” (FR: 1259). These concepts belonging more innately to doctrinal vocabulary or salvation history have comparably lower frequency ranks, and, similarly, concepts that connote hardship or strife are far away from the top of the list. This tendency is also found among the concepts more innately representing a Christian symbolic system, which we have listed: “heaven” is represented far more frequently than “hell”. “the holy spirit”, “forgiveness” “resurrection” and “mercy” are comparably more represented than “sin”, “judgement”, “satan” and “evil”. There is thus a clear preference in the corpus for a recognisable (i.e. not exclusively or distinct theological) vocabulary that in general exhibits a positive attitude; moreover, there is a clear preference for theological concepts that connote optimistic and hopeful perspectives on existence. This indicates a solicitude theme about caring for human life with God and Jesus at the centre.

The word count analyses are good examples of – in Moretti’s terms – a distant reading approach, as we read and analyse words outside of the immediate context of the texts they derive from. To some, these findings may come off as crude observations in an otherwise highly complex material. It can easily be argued that pastors may provide theological understandings of “resurrection”, “creation”, “sin”, “evil” and “salvation” without using these particular terms – it can be argued that these concepts are so abstract that it takes more words and complex sentence structures to adequately represent the content of those concepts. The word count analyses cannot account for the various contexts in which the words in the corpus can take part, but they can show us a clear preference for some terms over others, which adds

up to a seemingly coherent thematic framework. In the following, we will initially maintain a distant perspective on our text collection, but we will use tools that are more complex in order to get more intimate knowledge about the semantic structures in which the concepts are embedded throughout the sermon corpus.

Romance and Ethics

We have applied the seeded semantic network analysis to focus on the two theological concepts, *love* and *sin*, and their semantic embedding in the corpus.

In the primary associations to “love” in Figure 1, we find two word clusters. In one cluster, “mercy”, “forgiveness” and “grace” belong together. These terms all denote aspects of love that are God’s to give, and thereby indicate an implicit reference to God as loving and caring. The terms in the other cluster – “gentleness”, “munificence”, “devotion”, “unlimited” – contextualise love as something safe, kind and altruistic, but not through concepts necessarily associated with God. Overall, the terms in Figure 1 are mostly abstract terms represented through nouns, thus appearing as predicates to love. In comparison, the primary associations to “sin” in Figure 2 are nouns as well as verbs. This network is not as tightly connected as Figure 1, but there is a tendency towards nouns and verbs belonging to different clusters. We see that “sin” in particular shares the same contexts as “original sin” and “evil”, indicating that sin is a lasting and basic condition. The small cluster of yellow nodes inscribes sin in a context of movement through the terms “trespassing”, “sinner” and “owe”, signifying that sin is process-related and understood in terms of actions. The grey nodes constitute the most close-knit cluster of the three, which includes the terms “repent”, “atone”, “forgive” and “punishment”. Instead of contextualising sin in terms of direct actions, these words rather contextualise it in terms of *reactions* and frame sin as something that can be dealt with through deeds. Even though the network around “love” suggests different contexts, “love” overall appears as something lasting and positive in a rather close-knit network; there are no verbs indicating that love can be lost or is temporary. In comparison, “sin” seems to exhibit more nuances, as it at the one end of the spectrum is associated with stable conditions, while it at the other end is associated with deeds that can overcome or at least deal with sin. Sin thus appears as a more flexible concept in the corpus.

In Figure 3, however, we can interpret both seed terms through more complex structures. In this graph, we have included the secondary associations to the seed terms, and we may observe how the semantic fields around each seed term relate to each other. Even though some terms mostly related to “sin” also relate to “love”, there are no direct connections between “love” and “sin”. When we expand the associative structure around love, we find that the

altruistic conceptualisation of love becomes more consolidated through terms such as “unconditional”, “compassion”, “kindness”, “tenderness” and “to love”. Furthermore, with these types of terms, the altruistic love seems to denote a romanticised, intimate love. These terms dominate the top of the graph, where they are very closely related. The terms are moreover quite unambiguous – none of them indicate downsides or challenges associated with love, and none of them refer to love as wild or passionate. Instead, the romanticised representation of love seems rather uncomplicated, carefree and safe. Meanwhile, the conceptualisation of love as “God’s love” still seems to be in its own little cluster separate from the close-knit romanticised love and is thus less predominant in the semantic network around “love”. In the expanded network around “sin”, there are more connections between the terms than in Figure 2, but they are still not as closely connected as in the expanded network around “love”. However, the terms added to the network around “sin” accentuate the *action/reaction* aspect of sin rather than the *stability* aspect. Furthermore, with terms such as “revenge”, “regret”, “guilt” and “injustice”, the semantic field around sin is increasingly becoming an ethical domain.

We thus find semantic fields of core theological content through the semantic network analyses. However, we also see that the term “love” transgresses a delimited narrative of salvation history and instead takes part in a more mainstream understanding of love through a recognisable vocabulary. This insight supports our word count analyses, which illustrated that the traditional authoritative theological concepts are not necessarily used to represent the main themes in the sermon corpus. The representation of a romanticised love would emphasise the solicitude theme, where the close relations in the here-and-now seem enhanced. The semantic network analyses further support the finding that there is an overall positive attitude to be found in the corpus: love does not seem to have downsides, and the action/reaction pattern indicates that sin – while it may be associated with harsh conditions – can be coped with.

Thematic Allegiance to the Liturgical Year

Although our word count analyses and semantic network analysis of love show that pastors seem to present thematic content through a recognisable and relatable vocabulary, our topic modelling indicates that pastors also exhibit strong allegiance to the narrative structure of the liturgical year.

Table 2 shows that the terms “god”, “jesus”, “human beings”, “day” and “life” from the top of the MFW list (Table 1) appear in most of the eleven topics, which confirms that these terms are central to the thematic content in the corpus. Meanwhile, Table 2 also demonstrates that

these central terms take part in different word configurations in the topics. A large part of the vocabulary that unfolds in relation to the MFW terms in the topics refers to narrative content from biblical sources, which inscribes the MFW terms in different semantic contexts. Below, we characterise the thematic content that appears from these contexts in topics relating to holiday seasons. We focus specifically on Christmas, Easter, Pentecost and the first part of the Trinity season in order to observe variations throughout the course of the liturgical year.

The topic most related to the sermons from Christmas and New Year [CHRIST-NEW] clearly exhibits a Christmas theme, which is emphasised through the term itself, “christmas”. The nativity motif of the gospels is also distinctly apparent: “god”, “child”, “born”, “the child”, “bethlehem” and “joseph”. Jesus is not at the centre of this topic with references to “jesus”, “christ”, “saviour” or “messiah”, but as a child. His role in the Christmas sermons is thus framed in terms of his human nature and not his transcendent. A large part of the remaining terms in the topic are high on the MFW list: “world”, “year”, “light”, “human beings” and “joy”. Along with the topic-specific term “peace”, the message at the centre of the Christmas theme seems to be a wish for world harmony – a framing that resembles the optimistic attitude found in the solicitude theme from our word count analyses. However, here it unfolds in a clear context of Jesus entering the world at Christmas.

In the topic related to the sermons from the Easter tide [EASTER], we also see a strong allegiance to the holiday theme. We find in this topic that a few terms from the MFW list – “jesus”, “death/dead”, “the death”, “life” – interrelate with a larger vocabulary that clearly represents the Easter narratives as found in the gospels: “resurrection”, “jerusalem”, “died”, “the tomb”, “the disciples”, “easter morning”, “good friday” and “the cross”. We thus see that the thematic content in Easter sermons resonates less with the larger solicitude framework of the collection than the Christmas topic does. We also find that a more explicit theological vocabulary, as opposed to the solicitude vocabulary, permeates the Easter topic. For example, “resurrection” emerges as a distinctly central term in this topic, while it is absent in the overall solicitude theme. This observation indicates that the occurrence of the term is closely linked with its biblical narrative context, whereas terms such as “light” and “joy” are significant for, but not limited to, the nativity motif.

“the holy spirit” seems to follow a similar pattern as “resurrection” in the liturgical year, as it only appears to be central in the topic related to sermons following the Easter tide [POST-EAST_2]. We have previously described this sub-group of sermons as a bit artificial, as the holidays in this group are not as strongly related as for example the Easter sermons [EASTER]. However, through the terms in topic 8 in Table 3 [POST-EAST_2], we see a distinct Pentecost narrative: The Holy Spirit entering the world into the lives of human beings after Christ’s

ascension to heaven. Yet outside the contexts of Easter and Pentecost in the liturgical year, the terms “resurrection” and “holy spirit” are not emphasised as central theological components. Instead, they mostly seem to appear in the corpus in terms of their biblical narrative contexts. In the ELCD, the gospel pericopes of the Trinity season include motifs emphasising what a righteous life of Christian faith entails. Many of these in the form of parables, disseminating the teachings of Jesus. These stories are, among others, the narratives of the prodigal son (Luke 15:11–32), the camel and the eye of a needle (Matt 19:16–26), the house and the storm (Matt 7:22–29), the feast of the rich man (Luke 14:16–24), and the conversion of Zacchaeus (Luke 19:1–10). From our topic modelling, we found two different topics most related to sermons from the first part of the Trinity season, and they seem to emphasise the motif of righteousness and faith from different perspectives. In one topic, topic 10, there is a focus on ethical demands with a dichotomy of right and wrong. This framing includes active verbs and shares similarities with the dynamics in the semantic modelling of “sin”. The other topic, topic 3, does not seem to emphasise judgment and guilt, but rather a message of God’s love for human beings – especially the story of the prodigal son appears to belong to this context. In this topic, the semantic context of “love” promotes an understanding of care for human life, but it is not constructed as a romanticised love similar to the construction found in the semantic network analysis. Rather, “love” seems to be associated with the biblical sources, thus behaving similarly to the terms “resurrection” and “holy spirit”. Furthermore, in topic 3, Peter appears as a central figure. In the pericopes from this part of the Trinity season, Jesus’ appointment of Peter as the rock of faith (Matt 16:13–26) and Jesus’ proclamation of Peter as a fisherman appointed to catch people instead of fish (Luke 5:1–11) occur. Therefore, there is also a focus on the foundation of the Christian congregations, which is rooted in the stories of Peter.

The topic modelling analysis broadens our understanding of how theological concepts and themes are represented in the sermons. When we look at themes in relation to the church year, we find distinct theological discourses that unfold in close relation to the biblical narratives that structure the church year. So even though this discursive content is not correspondingly apparent from our MFW list, it does not mean that traditional authoritative theological concepts are not pertinent in this corpus. Rather, these concepts seem to unfold and become particularly central in expounding the biblical narratives.

7. Discussion: Theological Dynamics in Sermons

The three approaches applied unveil different thematic aspects of the sermon corpus. This unveiling informs us that the corpus covers a complex semantic field of diverse dynamics. In these dynamics, the role of theological discourse varies. In Woodhead’s explanation, theology

is defined as an academic discipline where a confessional basis and doctrinal discourse prevail, as she associates theological expertise with a power dynamic from above. If this dynamic is devalued, theological displacements occur. According to this terminology, the solicitude theme from the larger thematic framework of the sermons indicates at first sight that innate theological concepts are displaced. Here, a doctrinal vocabulary seems to be downplayed in favour of a vocabulary emphasising existential themes and care for human life. This framing points to an orientation towards inter-personal relations. These findings would support sociologist Robert Bellah’s early characterisation of symbolisation in modern Christianity. In 1964, he described the symbolic universe as a flexible and dynamic system that adapted to the situation of modern individualised human beings, where its transcendent and dualistic aspects would not be central anymore. Instead, there would be an emphasis on the human situation here and now³⁵. However, when we turn to our topic modelling results, we find a remarkable theological structure below the most frequently used words. We see that many of the terms on the MFW list are part of topics closely related to the liturgical year. The topics distinctly represent biblical narratives, as they very precisely combine agents, locations, liturgical tide and concepts of narrative significance. When the sermon corpus was built, we made sure to remove any representation of gospel pericopes in full text from the sermon manuscripts in cases where pastors had included the passage they would read aloud immediately before giving their sermon³⁶. Therefore, the biblical content found in the analyses is paraphrased by the pastors themselves and not copy-pasted directly from the gospel pericopes to the sermon manuscripts. Furthermore, even though we have pre-structured the corpus for the topic modelling analyses by concatenating the sermons, the design allowed the analyses to show an arbitrary relationship between sermons, if there had in fact not been a strong affiliation between sermons from the same liturgical seasons. Thus, pastors are in general very loyal to the theological premises of the church year.

Meanwhile, the seeded semantic network analysis demonstrates that an innate theological concept such as *love* does not belong solely to a doctrinal discourse, but is also embedded in a more popular cultural idea of love. In Nielsen and Johansen’s study of new Valentine’s Day services in the ELCD, they found that parishes tend to frame this event as a celebration of romantic love in couples’ lives³⁷. Combined, these observations indicate that a concept such as

³⁵ Robert N. Bellah, Religious Evolution, in: American Sociological Review 29.3 (1964), 358–374.

³⁶ Anne Agersnap, Ross Deans Kristensen-McLachlan et al., Sermons as data: Introducing a corpus of 11,955 Danish sermons, in: Journal of Cultural Analytics 12 (2020), 1–27.

³⁷ Marie Vejrup Nielsen & Kirstine Helboe Johansen, Transforming churches: the lived religion of religious organizations in a contemporary context, in: Journal of Contemporary Religion 34 (2019), 509–527.

love provides a link between an original theological context and a broader mainstream cultural awareness – the concept seems to have an ability to be transposed to a new context. However, not all components in the symbolic system seem suitable for such a linking according to our analyses. “sin” does not provide a direct link between cultural contexts, but it does display a flexibility between semantic contexts, and it exhibits an optimism, where sin as an alterable condition resembles the overall positive and hopeful attitude found in the MFW list. Even more so than with “sin”, we are able to observe a remarkable elasticity in the semantic universe of sermons through the term “love”, as it is part of thematic contexts in all three analyses. It is one of the most frequently used terms in the corpus, thus providing content to the solicitude theme. This motif of inter-personal relationships is even more unfolded through the term’s embedding in the semantic network analysis. However, it also obtains a liturgical pertinence in the first part of the Trinity season. Here, love is not romanticised, but seems to denote God’s love for human beings similarly to a father’s love for his son as told through parables.

The analyses closely examining terms and vocabulary challenge a distinction where theological discourse per se belongs to a dynamic from above strictly bound to doctrinal teachings. Rather than something only from above, the analyses indicate that theological terminology can also interact with other cultural codes. In the sermons, theological discourse can be observed from above as well as from below – the latter as a discourse where theological content is the outcome of negotiations with contemporary culture and the everyday lives of churchgoers. Therefore, even though this dynamic might be deemed a displacement of theology in Woodhead’s terminology, there would be no grounds for claiming that sermons are no longer a distinct domain of theological expertise. Rather, different dynamics are prevalent simultaneously in the sermons, which indicates that sermons are compounded phenomena where theological discourse is emergent – whether in relation to the church year or in relation to contemporary cultural codes.

8. Conclusion

The distant reading of sermons unveils collective patterns. It is a method that shows tendencies in the corpus and, in our approach, dynamics that can be interpreted closely. However, our study also demonstrates the value and necessity in applying different computational tools in order to observe the dynamics that occur simultaneously in the corpus – relying on just one of the three tools would not have made the thematic complexity in the corpus obvious. Since we study the collection as a collective text production, it has not been our intention to look for individual variations in the corpus. However, in the topic modelling design we did contrast the collective signal with a pastor signal, as we labelled sermons

according to holiday name as well as pastor pseudonym. Our finding that sermons thematically cluster more according to holiday tides than personal traits show the potential in studying sermons as a collective text system. Meanwhile, just as we do not study individual variations, nor do we claim an average way of composing sermons. In this way, the collective perspective allows us to infer tendencies in sermons based on some core thematic structures in the corpus without looking at individual cases or claiming a uniform genre. The structures are dynamics that occur simultaneously without precluding one and other³⁸.

Our approach to sermons allows us to address an aspect of preaching different from the oral event. We do not know for sure that the sermons in the corpus have been uttered in church precisely as they were written, and we have no insights into pastors’ individual performative styles. Instead, we focus on the fact that sermon documents exist and that they unfold an interpretive practice. Therefore, even though our approach does not analyse the oral communicative context of sermons, we are still interested in how content is communicated – however, not in terms of individual style, but in terms of implicit dynamics in the collective production of sermons. When we call sermons a collective testimony to Christianity and time, we thus allude to an understanding of the documents as a theological reality that only reveals itself as a collective phenomenon. Sermons are an ongoing documentation of Christianity and history, where pastors bear witness to tradition and the cultural context in which they expound tradition.

Anne Agersnap, PhD., born 1990, is currently a research assistant at the Centre for Grundtvig Studies at Aarhus University.

E-mail: anag@cas.au.dk

Kristoffer L. Nielbo, PhD., born 1977, is associate professor and director of Centre for Humanities Computing Aarhus at Aarhus University.

E-mail: kln@cas.au.dk

³⁸ Anne Agersnap, *Collective Testimonies to Christianity and Time*, PhD dissertation, Aarhus University 2021, 160–161.

Appendix A

The dictionary below include the English terms as represented in Figure 1, Figure 2 and Figure 3 and the original terms from the sermon corpus in Danish. The translations from Danish to English are made by the authors.

ENGLISH	DANISH
atone	sone
atonement	soning
compassion	medfølelse
confidence	tiltro
cruelty	grusomhed
devotion	hengivenhed
evil	ondskab
forgive	tilgive
forgiveness	tilgivelse
generosity	generøsitet
gentleness	mildhed
goodness	godhed
grace	nåde
guilt	skyld
hardening	forhærdelse
inclusiveness	rummelighed
injustice	uretfærdighed
kindness	venlighed
love	kærlighed
(to) love	elske
mercy	barmhjertighed
munificence	gavmildhed
original sin	arvesynd
owe	skylde
propitiatory sacrifice	sonoffer
publican	tolder
punishment	straf
regret	fortryde

repent	angre
revenge	hævne
sin	synd
sinner	synder
stupidity	dumhed
tenderness	ømhed
tolerance	overbærenhed
trespassing	overtrædelse
trust	tillid
unconditional	betingelsesløs
unlimited	grænseløs
wholehearted	ubetinget

Appendix B

The table below illustrates Danish liturgical holidays assigned to sub-groups. The table includes name of holiday sub-group, concrete holiday and name of holidays from the corpus metadata. Sub-groups are named based on Danish holiday names. Number of holidays in the Epiphany tide and Trinity tide varies depending on when Easter falls in a given year, so there are not always respectively seven and twenty six holidays in the two tides.

Holiday sub-group	Holiday (English)	Holiday (Danish)	Metadata annotation
ADVENT	1st Sunday of Advent	1. søndag i advent	% 1Adv
	2nd Sunday of Advent	2. søndag i advent	% 2Adv
	3rd Sunday of Advent	3. søndag i advent	% 3Adv
	4th Sunday of Advent	4. søndag i advent	% 4Adv
CHRIST-NEW	Christmas Eve	Juleaften	% Jul
	Christmas Day	Juledag	% JulD
	Saint Stephen's Day/2nd Day of Christmas	Sankt Stefansdag/2. Juledag	% 2Jul
	Christmas Sunday	Julesøndag	% JulS
	New Years Day	Nytårsdag	% Ny
	New Years Eve	Nytårsaftensdag	% NyA\SÆR
EPIPHANY	Epiphany Sunday	Helligtrekongers søndag	% Hel
	1st Sunday after Epiphany	1. søndag efter Helligtrekonger	% 1Hel
	2nd Sunday after Epiphany	2. søndag efter Helligtrekonger	% 2Hel
	3rd Sunday after Epiphany	3. søndag efter Helligtrekonger	% 3Hel
	4th Sunday after Epiphany	4. søndag efter Helligtrekonger	% 4Hel
	5th Sunday after Epiphany	5. søndag efter Helligtrekonger	% 5Hel
	Last Sunday after Epiphany	Sidste søndag efter Helligtrekonger	% SHel
LENT	Septuagesima Sunday	Søndag Septuagesima	% Sept
	Sexagesima Sunday	Søndag Seksegesima	% Seks
	Lent	Fastelavns søndag	% Fas
	1st Sunday of Lent	1. søndag i Fasten	% 1Fas
	2nd Sunday of Lent	2. søndag i Fasten	% 2Fas
	3rd Sunday of Lent	3. søndag i Fasten	% 3Fas
	Laetare Sunday	Midfaste søndag	% MidF
	Feast of the Annunciation	Mariæ Bebudelsesdag	% MBeb
EASTER	Palm Sunday	Palmesøndag	% Pal
	Maundy Thursday	Skærtorsdag	% Skær
	Good Friday	Langfredag	% Lang
	Saturday before Easter	Påskelørdag	% PåsL
	Easter Day	Påskedag	% PåsD
	2nd Day of Easter	2. Påskedag	% 2PåsD

POST-EAST_1	1st Sunday after Easter	1. søndag efter Påske	% 1Pås
	2nd Sunday after Easter	2. søndag efter Påske	% 2Pås
	3rd Sunday after Easter	3. søndag efter Påske	% 3Pås
	Common Prayer Day	Bededag	% Bed
POST-EAST_2	4th Sunday after Easter	4. søndag efter Påske	% 4Pås
	5th Sunday after Easter	5. søndag efter Påske	% 5Pås
	Christ’s Ascension	Kristi Himmelfartsdag	% KHim
	6th Sunday after Easter	6. søndag efter Påske	% 6Pås
TRINITY_1	Pentecost	Pinsedag	% Pin
	2nd day of Pentecost	2. Pinsedag	% 2Pin
	Trinity Sunday	Trinitatis søndag	% Trin
	1st Sunday after Trinity	1. søndag efter Trinitatis	% 1Trin
TRINITY_2	2nd Sunday after Trinity	2. søndag efter Trinitatis	% 2Trin
	3rd Sunday after Trinity	3. søndag efter Trinitatis	% 3Trin
	4th Sunday after Trinity	4. søndag efter Trinitatis	% 4Trin
	5th Sunday after Trinity	5. søndag efter Trinitatis	% 5Trin
TRINITY_3	6th Sunday after Trinity	6. søndag efter Trinitatis	% 6Trin
	7th Sunday after Trinity	7. søndag efter Trinitatis	% 7Trin
	8th Sunday after Trinity	8. søndag efter Trinitatis	% 8Trin
	9th Sunday after Trinity	9. søndag efter Trinitatis	% 9Trin
TRINITY_2	10th Sunday after Trinity	10. søndag efter Trinitatis	% 10Trin
	11th Sunday after Trinity	11. søndag efter Trinitatis	% 11Trin
	12th Sunday after Trinity	12. søndag efter Trinitatis	% 12Trin
	13th Sunday after Trinity	13. søndag efter Trinitatis	% 13Trin
TRINITY_3	14th Sunday after Trinity	14. søndag efter Trinitatis	% 14Trin
	15th Sunday after Trinity	15. søndag efter Trinitatis	% 15Trin
	16th Sunday after Trinity	16. søndag efter Trinitatis	% 16Trin
	17th Sunday after Trinity	17. søndag efter Trinitatis	% 17Trin
TRINITY_3	18th Sunday after Trinity	18. søndag efter Trinitatis	% 18Trin
	19th Sunday after Trinity	19. søndag efter Trinitatis	% 19Trin
	20th Sunday after Trinity	20. søndag efter Trinitatis	% 20Trin
	All Hallow’s Day	Allehelgensdag	% AHelg
TRINITY_3	21st Sunday after Trinity	21. søndag efter Trinitatis	% 21Trin
	22nd Sunday after Trinity	22. søndag efter Trinitatis	% 22Trin
	23rd Sunday after Trinity	23. søndag efter Trinitatis	% 23Trin
	24th Sunday after Trinity	24. søndag efter Trinitatis.	% 24Trin
TRINITY_3	25th Sunday after Trinity	25. søndag efter Trinitatis.	% 25Trin
	Last Sunday of the liturgical year	Sidste søndag i kirkeåret	% Sidst

Appendix C

An excerpt of the 20 topics in Danish represented with the top seven words in each topic.

	Word 0	Word 1	Word 2	Word 3	Word 4	Word 5	Word 6
Topic 0	gud	jesus	liv	dag	livet	tro	verden
Topic 1	jesus	gud	guds	tro	maria	rige	dag
Topic 2	kærlighed	kristus	verden	kristne	døden	blot	kirken
Topic 3	jesus	gud	rige	guds	liv	mennesker	peter
Topic 4	jesus	peter	gud	hyrde	liv	thomas	sandheden
Topic 5	findes	magt	onde	fast	hverken	står	altså
Topic 6	jesus	død	jesu	døden	opstandelse	liv	jerusalem
Topic 7	liv	livet	brød	mske	leve	gud	guds
Topic 8	gud	guds	jesus	ord	jesu	kristus	helligånden
Topic 9	små	gamle	mænd	børn	vand	sted	barn
Topic 10	siger	ret	fået	skyld	sagde	plads	gå
Topic 11	jesus	gud	liv	ord	guds	martha	dag
Topic 12	jesus	mennesker	gud	altså	år	kristne	folk
Topic 13	gud	ord	menneske	vej	guds	netop	livet
Topic 14	dagens	jesus	altså	kærlighed	tekst	tror	guds
Topic 15	ord	lys	dag	siger	går	hele	står
Topic 16	gud	julen	barn	guds	verden	jul	år
Topic 17	johannes	jesus	dag	tid	lys	guds	advent
Topic 18	gud	jesus	guds	liv	livet	mennesker	tro
Topic 19	dét	menneske	frem	én	billede	fortællingen	sat