

THE USE OF DIACHRONIC CORPORA IN STUDYING MEDIA TEXTS

Teshabayeva Dilfuza Mo'minovna
f.f.d., professor
O'zDJTU

Annotation

This article explores the use of diachronic corpora in analyzing media texts, focusing on linguistic evolution, stylistic shifts, and discourse dynamics over time. It examines how lexical, thematic, and stylistic changes reflect broader societal and cultural shifts, highlighting implications for both media studies and linguistics. By integrating quantitative and qualitative methods, the article offers insights into how diachronic corpora contribute to understanding language change in media contexts.

Keywords

Diachronic corpora, media texts, linguistic evolution, discourse analysis, lexical change, stylistic shifts, media studies, corpus linguistics, language change, journalistic style.

The analysis of media texts using **diachronic corpora** has become an important research area within corpus linguistics, offering insights into how language, style, and communicative strategies in media have evolved over time. Media texts, such as newspapers, magazines, and digital publications, play a significant role in shaping public opinion and reflecting societal changes. By examining these texts through a diachronic lens, researchers can understand trends in language use, shifts in journalistic style, and changes in media discourse (Partington et al., 2020). This article explores the application of diachronic corpora to the study of media texts, emphasizing their impact on understanding linguistic evolution and media dynamics.

Diachronic corpora are collections of texts from different time periods, compiled to study language change over time. These corpora are particularly useful in analyzing media texts, which are often sensitive to cultural, political, and social shifts (Baker, 2006). Media texts, due to their broad dissemination and influential role in society, are ideal sources for diachronic studies, allowing researchers to track changes in language, rhetoric, and discourse over years or even decades.

A diachronic corpus contains texts from various historical periods, with consistent sampling methods to ensure representativeness (Sinclair, 1991). For example, the **SiBol Corpus** (Siena-Bologna Modern Diachronic Corpus) comprises UK newspaper texts sampled from 1993, 2005, and 2010, designed to maintain comparability while eliminating extraneous variables (Partington et al., 2020). This structure enables researchers to compare language use across different eras, focusing on factors like lexical choices, stylistic markers, and shifts in journalistic practices. The use of diachronic corpora has allowed researchers to examine how language in media texts has changed over time, reflecting broader sociolinguistic trends.

One of the most evident changes in media texts is in **lexical usage**, where new words emerge, old words fade, and the frequency of existing words shifts based on current events, technological advancements, or changing societal norms. For example, Schulz et al. (2020) used diachronic corpora to study the lexical patterns of political discourse in Austrian German newspapers over 20 years, revealing both the emergence of new political terminology and the decline of outdated terms.

Diachronic corpora are also instrumental in identifying **stylistic shifts** in media texts. Studies have shown a gradual movement towards more informal, conversational styles in journalistic writing over time (Claire, 2010). This trend aligns with the popularization of news media, where writers increasingly use simpler language, shorter sentences, and more direct communication to engage wider audiences (Biber et al., 1999). For instance, the analysis of British newspapers in the SiBol Corpus indicated a marked shift towards colloquial expressions, reflecting the evolving relationship between journalists and readers (Partington et al., 2020). The examination of diachronic corpora offers deeper insights into **discourse evolution** within media texts, highlighting how media outlets construct narratives and present information over time.

Subjectivity and stance, which refer to the degree of personal judgment and positioning within a text, are crucial elements of media discourse (van der Wouden et al., 2012). Diachronic corpora facilitate the analysis of how these elements have changed in response to social and political developments. For example, van der Wouden et al. (2012) found that Dutch news texts displayed increasing levels of subjectivity over time, reflecting a shift towards more opinionated journalism.

Thematic changes in media discourse can also be tracked through diachronic corpora. As media texts respond to ongoing social and political events, certain themes may become more prominent, while others may decline. For instance, Camiciottoli's (2021) study of **Vogue** magazine articles over several decades revealed shifts in thematic focus, particularly in fashion journalism, as influenced by changing cultural norms and trends. The corpus analysis showed that topics related to sustainability and body positivity gained more attention over time, mirroring broader societal trends.

The application of diachronic corpora in studying media texts requires a combination of quantitative and qualitative methods. These approaches ensure a comprehensive understanding of language change and media dynamics. The annotation of diachronic corpora is essential for identifying key linguistic features and discourse markers. Automated tools, such as part-of-speech (POS) taggers and named entity recognizers (NERs), are commonly used to annotate corpora, making it easier to track changes in specific linguistic elements (Baker, 2006). In media studies, annotating elements like modality, metaphor, and intertextual references is particularly useful for examining discourse evolution (Partington et al., 2020).

Researchers often employ **comparative** and **longitudinal** analyses when working with diachronic corpora. Comparative analysis involves examining differences between time-specific corpora to understand lexical, stylistic, and thematic changes (Claire, 2010). Longitudinal analysis, on the other hand, focuses on tracing the development of specific linguistic or thematic elements over extended periods, offering a more nuanced understanding of media discourse dynamics (Schulz et al., 2020). The use of diachronic corpora in studying media texts has significant implications for both media studies and linguistics, providing insights into how media language evolves in response to sociopolitical contexts and changing reader expectations.

From a linguistic perspective, diachronic corpora offer valuable data for understanding language change, particularly in media, where language adapts rapidly to societal shifts. This understanding not only contributes to corpus linguistics but also aids in the development of language models for natural language processing (NLP) applications, such as sentiment analysis and automated text summarization (Biber et al., 1999). For media studies, insights from diachronic corpus analysis can inform journalistic practices by highlighting how language choices impact audience engagement and perception over time. By understanding the historical trends in media language and style, journalists can craft more effective and context-sensitive narratives (Camiciottoli, 2021).

The study of media texts using diachronic corpora provides a comprehensive view of linguistic evolution, stylistic trends, and discourse dynamics. By analyzing lexical changes, stylistic shifts, and discourse evolution, researchers gain valuable insights into the ways in which media texts reflect and shape societal changes. Future research can expand on these findings by incorporating more diverse media sources, exploring the impact of digital media formats, and analyzing the effects of globalization on media language.

References

1. Baker, P. (2006). *Using corpora in discourse analysis*. Continuum.
2. Biber, D., Conrad, S., & Reppen, R. (1999). *Corpus linguistics: Investigating language structure and use*. Cambridge University Press.
3. Camiciottoli, B. C. (2021). Hyphenated phrasal expressions in fashion journalism: A diachronic corpus-assisted study of *Vogue* magazine. *Journal of Pragmatics*, 186, 24-36.
4. Claire, R. (2010). A diachronic analysis of shifting newspaper language style, 1993–2005. *Discourse Studies*, 12(4), 427-449.
5. Partington, A., Duguid, A., & Taylor, C. (2020). *Modern diachronic corpus-assisted discourse studies (MD-CADS) on UK newspapers*. John Benjamins.
6. Schulz, S., Strohmaier, M., & Borghi, A. (2020). Comparing lexical usage in political discourse across diachronic corpora. *Proceedings of the International Conference on Language Resources and Evaluation (LREC)*, 288-297.
7. Sinclair, J. (1991). *Corpus, concordance, collocation*. Oxford University Press.
8. van der Wouden, T., Foolen, A., & van den Berg, M. (2012). Diachronic changes in subjectivity and stance: A corpus linguistic study of Dutch news texts. *Language Sciences*, 34(5), 567-583.