

Barantsev Ya. A.,
PhD Student in Philology,
Assistant Lecturer at the Department of Romano-Germanic Philology and Professional Communication
Ternopil Volodymyr Hnatiuk National Pedagogical University
<https://orcid.org/0009-0000-4795-4598>

COMPARATIVE CONTENT ANALYSIS OF THE DOMINANT LINGUOCULTURAL CONCEPTS *LANGUE / LENGUA* AS REALISED WITHIN CONTEMPORARY FRENCH-LANGUAGE AND SPANISH-LANGUAGE MEDIA DISCOURSES (BASED UPON ARTICLES FROM FRENCH, CANADIAN, SPANISH, AND MEXICAN PERIODICALS)

Summary. The present article attempts to review and test the application of comparative content analysis of the dominant linguocultural concepts *LANGUE / LENGUA* as realised within contemporary French-language and Spanish-language media discourses based upon selected articles from the periodicals *Le Figaro* (France), *Le Devoir* (Canada), *El Mundo* (Spain), and *La Jornada* (Mexico). The research emphasises that the concept *LANGUAGE* is of foremost significance in the worldview of every nation. Hence, there arises the question of comparing the intrinsic features of its realisation across world languages, particularly French and Spanish, which are linked not only by their international status, but also by typological characteristics shared across all levels of the linguistic hierarchy. Thus, the study tests the assumption grounded in the general similarity of the dominant linguocultural concepts *LANGUE / LENGUA* as realised within Francophone and Hispanophone media discourses drawing on periodicals from several countries, with comparative content analysis being employed.

Specifically, the analysis was carried out using a suite of Python libraries that supported the preliminary stages of text processing, including NLTK, StopwordsISO, Stanza, KeyBERT, Pandas, and Matplotlib. In addition, the data obtained through these computer-assisted tools were further augmented by the results of the semantic field method – which was used to group words into thematic macroconceptual clusters – and of discourse analysis, which rendered it possible to identify and delineate implicit strategies emerging from the repetition of certain culturally marked key words. In the course of the research, it became clear that the dominant concepts *LANGUE / LENGUA* play a substantial role in the conceptual organisation, actualisation of cultural experience, and argumentation in the selected articles.

Upon an examination of the texts at the level of their conceptual network and discursive strategies employed, it was concluded that both the French- and Spanish-language media discourses share a common conception of the pivotal role of language in national self-identification as well as in the establishment of international relations. It proved characteristic of the texts published in Canadian and French periodicals that they address the questions of safeguarding the language and protecting cultural heritage, whereas in the Spanish-language periodicals the notion of further advancing linguistic co-operation is brought to the fore. At

the same time, within the Mexican media discourse, there was likewise discerned the theme of national self-identification, along with the delineation of its distinctive markers, which, in turn, creates a point of connection with the Canadian text. Furthermore, it was found out that culturally marked precedent-related phenomena strengthen the author's argumentation and the effect of the strategies implemented. Finally, the article outlines the prospects for further testing of the analysis and the need for its application on a wider material with the involvement of additional Python libraries and linguistic methods to increase the representativeness of the research results based upon the integration of both quantitative and qualitative approaches.

Key words: media linguistics, dominant linguocultural concept, content analysis, linguistic image of the world, Romance languages.

Problem Statement. The concept *LANGUAGE* pertains to the fundamental elements of a nation's worldview, which shapes its social outlook, ethnocultural selfhood, and collective memory. It reflects the code, the distinctive traits of the national mind, and the ways in which the surrounding reality is conceptualised, thereby evoking specific historical context that underpins the cultural continuity of generations. Moreover, inasmuch as it represents the distinctiveness of the sign system by whose very means its own functioning is, in turn, determined – that is to say, inasmuch as there obtains an interdependence between the phenomenon, its essence, and its nomen – there are grounds for regarding it as a cultural metaconcept, and, owing to its central place in the axiological and cognitive images of the world, all the more so as a dominant cultural concept of a given linguistic community. Yet the question of the manner in which it finds its realisation across languages appears to remain a matter of considerable debate, in particular at the level of discursive space and the lexical-semantic associative networks that arise within it. By the same token, no less problematic is the issue whether the said dominant linguocultural concept (this term shall henceforth be abbreviated as DLC) evokes identical associations among speakers of the same language who belong to distinct historically constituted cultural communities – a circumstance that, for its part, necessitates thoroughgoing research conducted in compliance with a comprehensive framework.

The above considerations, in our view, make it needful to undertake a comparative inquiry into how the various manifestations of the DLC LANGUAGE are realised across a number of international languages, notably French and Spanish. Such an investigation is particularly pertinent within the bounds of media discourse, wherein societal events and transformations are continuously reflected. In addition, media discourse is a focus for highlighting a wide variety of ideas, the expression of which is closely related to the involvement of communicative strategies and tactics, as well as to cultural implications, diverse forms of intertextuality, and precedent-related cultural phenomena, which are not infrequently grounded in knowledge of cultural facts. Meanwhile, the employment of both quantitative and qualitative approaches, such as would facilitate the elucidation of those relations, appears to enable the processing of textual data, their interpretation from the viewpoint of discourse studies, and drawing conclusions about the significance of particular concepts across cultures. Thus, the research question addressed in this study consists in determining the way in which the concepts LANGUE / LENGUA are realised within the French- and Spanish-language media discourses based upon the material from periodicals of several countries, as well as in ascertaining how this representation may be examined using comparative content analysis.

Literature Review. The issues of media linguistics and cultural linguistics within the bounds of the anthropological paradigm have stood at the centre of scholarly inquiry since the end of the 20th century. A significant contribution to the theoretical foundations of critical discourse analysis (CDA) was made by T. A. van Dijk (the study of news as discourse) [1], N. Fairclough (the theory of media discourse) [2], R. Fowler (the language of news viewed through the prism of discourse and ideology) [3], and other scholars. Further issues concerning cultural linguistics with regard to other disciplines were addressed in the works of G. B. Palmer (the general theory of cultural linguistics) [4], F. Sharifian (the notion of cultural conceptualisation and language) [5], M. Montgomery (the connection between language, media, and culture) [6], and those of other researches. The continued elaboration of these questions within the said fields of linguistics, considered through the optic of an interdisciplinary approach, is to be found in the studies of M. Burger [7], C. Taylor and D. del Fante [8], D. Nguyen et al. [9], B. Gallardo Paúls [10], and of other authors. The salient theoretical and practical achievements of modern linguocultural research are reflected in the work of A. Zahnitko and I. Bohdanova [11]. I. Saievysh offers a comprehensive methodology for identifying dominant linguocultural concepts [12]. O. Ianytska analyses linguocultural aspects of the analysis of the concept FAITH in English, Ukrainian, and French [13]. In the works edited by V. Wiegand and M. Mahlberg [14], E. Friginal and J. A. Hardy [15], computer-assisted corpus processing methods are presented for linguistic research.

Research objective. The objective of the study is to outline the methodological potential of comparative content analysis of the DLC LANGUE / LENGUA within contemporary French- and Spanish-language media discourses based upon the material drawn from the periodicals *Le Figaro* (France), *Le Devoir* (Canada), *El Mundo* (Spain), and *La Jornada* (Mexico). In order to attain the objective, the following tasks are formulated: (1) to develop a set of methods for analysing linguocultural concepts within the selected text material; (2) to conduct a quantitative analysis

of linguocultural concepts identified within the articles, to compare and to interpret the results obtained; (3) to interpret implicit markers of discursive strategies implemented on the basis of the DLC as attested in the articles from the mentioned media sources.

Results and discussion. For the purpose of the present study, we selected articles from the leading French-language periodicals *Le Figaro* (France) and *Le Devoir* (Canada), on the one hand, and the Spanish-language ones *El Mundo* (Spain) and *La Jornada* (Mexico), on the other. This choice is justified by the representativeness of the sources under consideration: all the aforesaid periodicals rank among the most influential daily newspapers with long-established traditions in their respective countries, and they reflect the stance of society as concerns cultural and social matters.

It is understood that contemporary linguistic methodology is characterised by a broad integration of research tools, of which a significant part is grounded in the achievements of computational linguistics. It is likewise no coincidence that the potential of comparative, structural, semasiological, communicative-functional, cognitive, and other linguistic methods should have been enriched through the active development of those instruments that have arisen as a result of the expanding capacities of information technologies and their software, particularly in the domain of natural language processing. That said, we regard the use of digital tools developed in the Python programming language as promising, given that their advantages for linguistic research have already been described in scientific practice (see in particular [16]). It may be argued that making use of this high-level language can enhance the validity of the results yielded by the text analysis whilst also providing greater precision in determining the status of concepts as dominant within a particular culture, which, need hardly be admitted, may not always be carried out in as unambiguous a manner as one might expect.

For the purpose of our analysis, we applied a suite of Python libraries that facilitated a full cycle of preliminary text processing. These included NLTK (used for the tokenisation of the selected text corpus), StopwordsISO (to exclude stopwords from consideration), Stanza (for the lemmatisation and part-of-speech differentiation), KeyBERT (to isolate key concepts in processed texts), Pandas (to structure data in tables), and Matplotlib (to construct frequency diagrams of lexical units). The data obtained through these computer-assisted tools were further supported by the integration of methods drawn from semantic field analysis (employed to group words by dominant thematic categories) and discourse analysis. This combination was configured such that implicit strategies implemented through the repetition of individual culturally marked key words could be identified and subsequently described.

The initial stage of our investigation consisted in selecting the linguistic material and processing it by means of the methods outlined above. The primary criterion for choosing the texts was their shared topical focus, to wit, the role of language within the cultural dimension. With a view to conducting content analysis of French media discourse, we opted for the article *A quoi ressemblera la langue française de demain ?* dated 21 March 2024 (*Le Figaro*, France) [17]. From genre perspective, the article is an interview on the challenges facing the modern *francophonie*, the issue of multilingualism, and the future of the French language. As part of the study, the text of the article was analysed using automated language processing tools, which made it possible to

obtain structured data on frequency characteristics and dominant concepts that reflect the specifics of French-language discourse within a selected text segment. The token frequency analysis function was set to a sampling parameter of '15'. The following is a diagram displaying the most frequent culturally relevant tokens identified in the analysis (see Fig. 1):

As may be observed from the diagram, the most frequent tokens were found to be 'langue' (44), 'français' (31), 'francophonie' (8), 'francophone' (8), 'international' (6) 'maternel', (5), 'monde' (5), 'organisation' (4), 'visite' (3), 'locuteur' (3), 'notion' (3), 'réalité' (3), 'projet' (3), 'peuple' (3), and 'institution' (3), among which the word *langue* functions as the archiseme, while the remaining units correspond thereto on the basis of paradigmatic or syntagmatic relations, interacting with the core constituent of the lexical-semantic field through close associative links. To verify these results, we used the KeyBERT semantic model to identify the key concepts (i.e. semantic dominants) of the selected text – those with the greatest semantic proximity to its overall content. Taking lemmatisation into account, the following concepts turned out to be the most relevant: 'français', 'étatique', 'francosphère', 'francophone', 'international', 'France', 'québécois', 'linguistique', 'anglais' (see Table 1).

As a result, the quantitative analysis of the article allowed us to identify quite a number of key concepts that constitute the semantic core of the text. The most frequent lemma – 'langue' – appears as a universal linguocultural identity marker. In this context, the corresponding conceptual exponent *langue* is realised through the tactic of universalisation, which forms part

of the broader strategy of cultural legitimisation. Thus, language is presented as a fundamental social bond, capable of structuring experience and fostering social unity: *La langue est une manière de produire du sens, de s'orienter dans l'opacité du monde. De tous les liens que les hommes peuvent nouer dans une société, celui de la langue est peut-être le plus fort. Il en est de même sur le plan international : le fait de parler la même langue d'un pays à l'autre crée un lien particulièrement fort* [17]. The same word frequently appears in collocation with the token *française*; for instance, in the following fragment it is accompanied by statistical data, which, in turn, can be interpreted as a communicative tactic of quantitative confirmation and support for the fact that French holds the status of an international language: *320 millions, c'est le chiffre qu'avance l'Observatoire de la langue française, mis en place au sein de l'Organisation internationale de la Francophonie (OIF), et qui effectue une enquête en profondeur tous les deux ans* [17]. The lexeme *francophonie* plays an important role in the French-speaking worldview, and it also proved to be significant both in terms of frequency characteristics and semantic relevance: *La francophonie est à la fois une réalité et un projet. La réalité, c'est l'ensemble des « parlants français » à travers le monde. Le partage d'une même langue par des peuples qui n'appartiennent pas nécessairement à une même organisation étatique ou politique* [17]. It is noteworthy that in this context *langue* emerges not only as a social reality but also as a political and cultural project grounded in institutional structures ('la "francophonie" avec un grand "F" et un petit "f"'). Other significant concepts turned out to be *langue internationale, langue maternelle, organisation,*

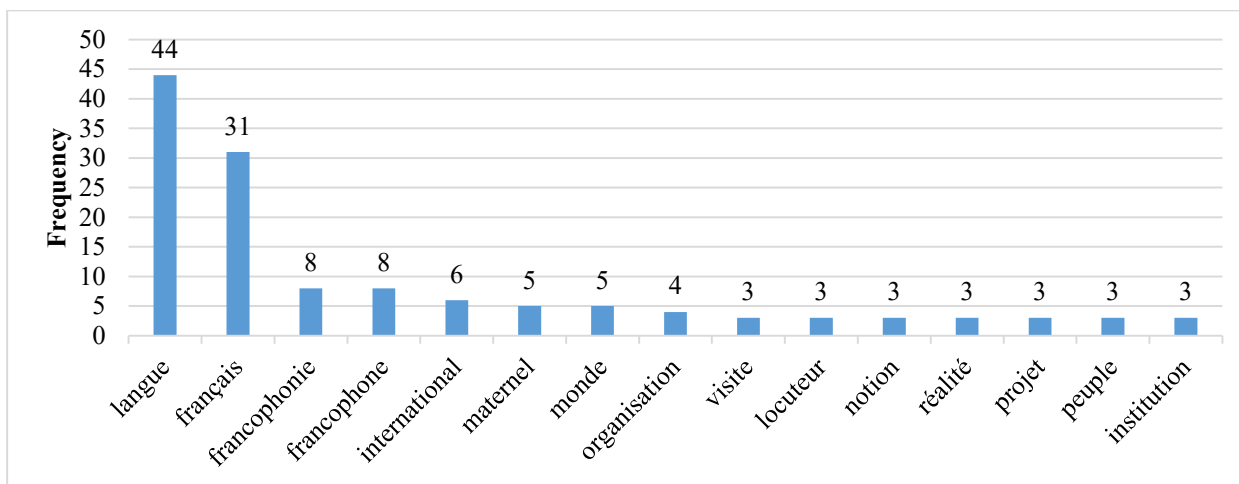


Fig. 1. 15 most frequent tokens in the text (the article from *Le Figaro*)

Table 1

Selected dominant concepts identified in the article from *Le Figaro*

No.	Key concept	Relevance index
1	'française'	0.4946
2	'français'	0.4341
3	'étatique'	0.3879
4	'francosphère'	0.3752
5	'francophone'	0.3717
6	'francophones'	0.3691
7	'internationales'	0.3688
8	'France'	0.3567
9	'québécois'	0.3494
10	'linguistiques'	0.3491

peuple, and France, whose appearance, as one may argue, realises a set of strategies, namely those of global coexistence, demographic projection, institutionalisation, cultural identification, and cultural continuity.

Another semantically dominant token that ought to be considered is *étatique*, which in its substantive form (*État*) recreates an associative connection with the state and its role in the establishment, protection, and development of the language. Although it is mentioned only once in the article itself, the automated processing algorithm positioned it as the third most relevant concept. This may be explained by the fact that it reveals an associative connection of content not merely with the cultural and educational but also with the political and institutional dimensions of discourse: the role of language is thus understood as an element of state policy, which ensures its international status, and therefore is closely connected with it: *Le partage d'une même langue par des peuples qui n'appartiennent pas nécessairement à une même organisation étatique ou politique* [17].

It is worth remarking that the analysis revealed links between the dominant concept LANGUE and precedent-related cultural phenomena identified in the text, viz., precedent-related names (*Ronsard, du Bellay*) and an allusion to a precedent-related text (*défendre et illustrer la langue française*, which refers to the literary theory text by the French poet J. du Bellay *La Défense et illustration de la langue française*). One may also consider as an example of a precedent-related phenomenon the toponym *Villers-Cotterêts*, which is known for the fact that King Francis I of France signed into law the Ordinance of Villers-Cotterêts (1539) there. Moreover, it is

known to be the birthplace of the writer Alexandre Dumas, père; nowadays, *la Cité internationale de la langue française* is located in this city.

With the object of studying the Canadian French-language media discourse, we selected the article *Célébrons notre langue française, celle qui nous unit et nous définit* of 18 March 2023 (*Le Devoir*, Canada) [18]. In terms of genre, it constitutes a statement issued by the ministers of the Government of Quebec on the eve of celebrating the International Day of Francophonie. The article highlights the role of the French language as the basis of the region's identity, as well as its continued presence in the international linguistic context. The most frequently occurring tokens identified through the analysis of this text are displayed in Figure 2.

To elaborate on these results, we conducted an analysis of the key concepts in the selected text; the most frequent of these, taking lemmatisation into account, are 'québécois', 'Québec', 'français', 'francophonie', 'national', 'devoir', 'sauvegarde', 'culturel', 'langue' (see Table 2).

It may be inferred from the statistical data presented that the most frequent key concepts are realised in connection with the application of certain communicative tactics and discursive strategies. The first semantically dominant concept identified in the article is the lexeme *Québec*, which carries distinct geographical and cultural marking. Quebec appears in the text as the centre of Francophonie on the American continent. Here, the tactic of self-identification is carried into effect as part of a broader strategy of cultural consolidation: *Le Québec est une nation dynamique, créative et unique au monde. Nous sommes le vaisseau amiral de la*

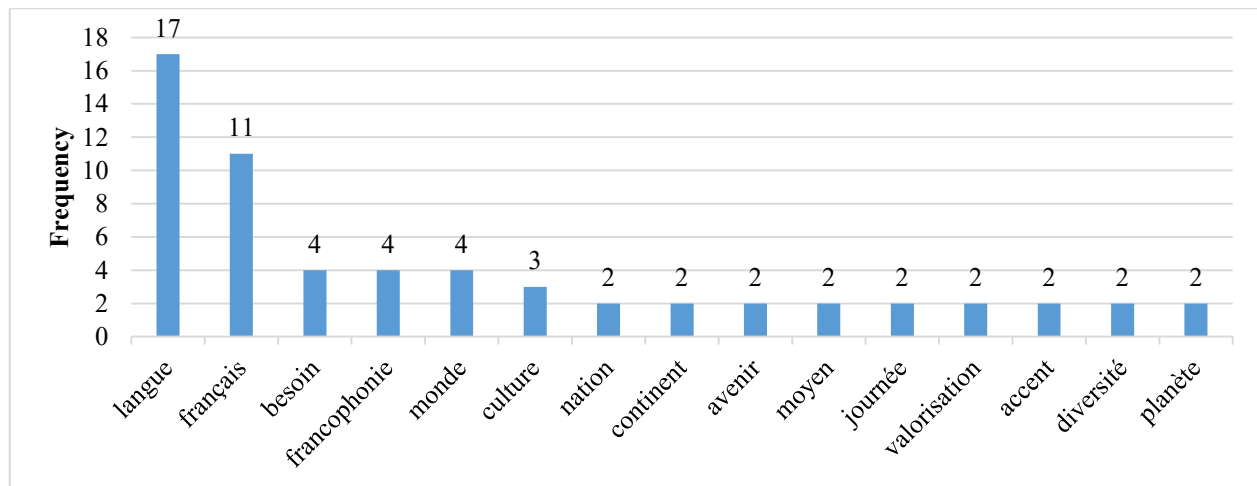


Fig. 2. 15 most frequent tokens in the text (the article from *Le Devoir*)

Table 2

Selected dominant concepts identified in the article from *Le Devoir*

No.	Key concept	Relevance index
1	'québécois'	0.4889
2	'Québec'	0.4466
3	'français'	0.4358
4	'française'	0.4318
5	'francophonie'	0.3247
6	'nationale'	0.3188
7	'devoir'	0.3125
8	'vaisseau'	0.3046
9	'sauvegarde'	0.2919
10	'statistique'	0.2888

langue française en Amérique. C'est elle qui nous distingue du reste du continent et du reste du monde [18]. As in the article from *Le Figaro*, and partly due to its topic, the token *français*, functioning both as an attributive marker and a substantivised concept, holds a central position in the text. It is noteworthy that here, too, one may observe the use of the tactic of quantitative corroboration intended to highlight the global presence of the French language: *Le français est la cinquième langue mondiale. Il est parlé par plus de 320 millions de personnes dans le monde. C'est également la troisième langue des affaires sur la planète* [18].

No less significant are the dominant concepts nominated by the lexemes *national* (from *nation*), *devoir*, *sauvegarde*, and *culturel* (from *culture*), which may be interpreted as being associated with several strategies at once: emphasising the state dimension of language policy, institutionalisation, cultural continuity, problematisation (in the sense of drawing attention to current issues of promoting the French language), protection, and cultural legitimisation: *Il est de notre devoir historique de prolonger cette présence linguistique; Mais la sauvegarde et la valorisation de la langue française ne sont pas uniquement des responsabilités du gouvernement; Car à l'image de la diversité écologique qui se doit d'être préservée sur notre planète, la diversité culturelle est une richesse pour tous les humains* [18]. The above tactics and strategies together determine that the media discourse of the article from *Le Devoir* positions the French language as part of the national and world cultural heritage that must be preserved for the future.

As in the previous case, the text attests to the employment of precedent-related cultural phenomena and allusions, whose role may be understood through their close association with the dominant concept *LANGUE*. For instance, considering the sentence *Nous sommes le vaisseau amiral de la langue française en Amérique* [18], one can identify the metaphorical expression *le vaisseau amiral*, which may be regarded as an allusion to a precedent-related situation, namely that of maritime history of Quebec since the expeditions of Jacques Cartier and Samuel de Champlain; it is telling that the flag of Quebec City depicts a ship.

In developing the analysis further, it should be noted that the concept *LENGUA* proved to function within Spanish-language media discourse in a broadly similar manner to the previously examined dominant concept *LANGUE*. In order to investigate its representation within the Hispanophone media-discursive space, we

chose the article *Don Felipe inaugura el Congreso de la Lengua: "El siglo XXI debe ser el siglo del español"* of 27 March 2023 (*El Mundo*, Spain) [19]. The topic of the article is the opening of the IX International Congress of the Spanish Language that took place that same year in Cádiz, during which the current state and prospects for the future development of the Spanish language were discussed. Although the article is primarily informative, it contains a number of key concepts thematically related to the lexical-semantic field 'Lengua', which are realised within the informative-communicative model defined by its intentionality and discursive strategies. For automated processing of the text, we applied the same sampling parameter to the 15 most frequent tokens (see Fig. 3).

Among the lexemes denoting the key concepts of the text and showing the greatest semantic proximity to its overall content, the relevant units, after lemmatisation, were as follows: 'español', 'nación', 'relación', 'hispanica', 'seguridad', 'constitución', 'felicidad', 'institucional', 'esencia', 'celebración', 'identidad' (see Table 3).

Further quantitative analysis of the article rendered it possible to identify a number of key words that form the semantic core of the text, the first of which being *español* (in its substantivised form), as in the following sentence (a passage from a speech delivered by King Felipe VI of Spain) about the future development of the Spanish language: *«El siglo XXI debe ser el siglo del español», dijo Don Felipe, que recordó que «en 2100 se prevé que el 6,3% de la población mundial podrá comunicarse en español»* [19]. The given fragment illustrates the implementation of the strategy of cultural legitimisation and demographic forecasting realised through the tactic of quantitative confirmation, likewise identified in the French-language examples. In the Spanish text, however, this tactic is distinctly future-oriented.

Given the topic of the article, the lexeme *lengua* holds the central place in the analysed text: *El español desde sus orígenes es una lengua mestiza y ese mestizaje trasciende a todas sus cultura, está en todas las naciones que lo hablan* [19]. Nevertheless, it is relevant to observe that this lexeme co-occurs in the text with its allolex *idioma*, as well as in the adjectival collocation *la lengua española: Lengua española, mestizaje e interculturalidad. Historia y futuro, es el lema del congreso y se puede explicar, sencillamente, como un intento de comprender el idioma español en relación a las lenguas y las culturas con las que convive, ya sean las otras lenguas*

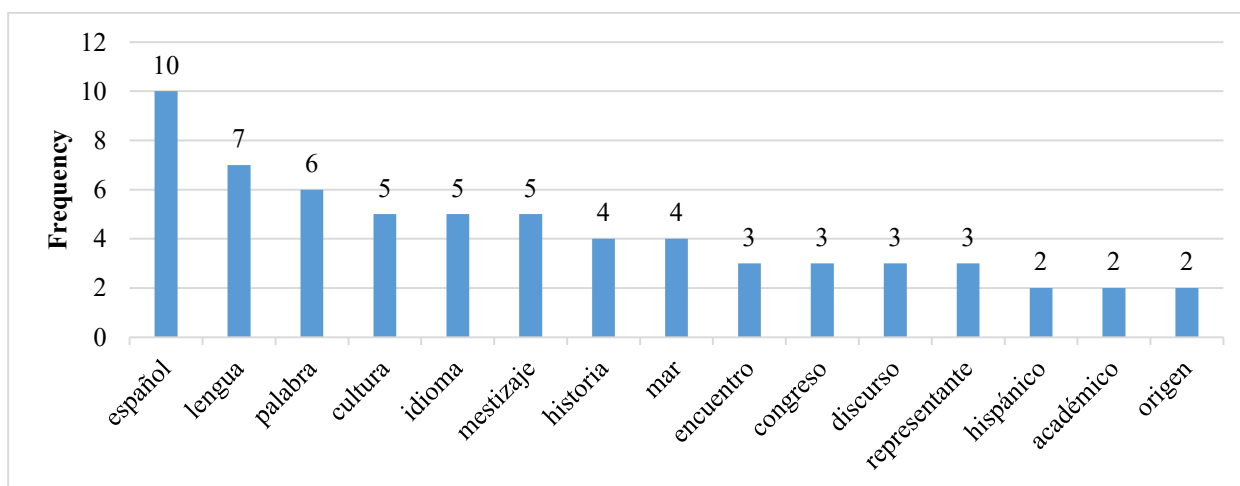


Fig. 3. 15 most frequent tokens in the text (the article from *El Mundo*)

Selected dominant concepts identified in the article from *El Mundo*

No.	Key concept	Relevance index
1	'español'	0.5022
2	'naciones'	0.4883
3	'relación'	0.4784
4	'hispanica'	0.4693
5	'seguridad'	0.4639
6	'constitución'	0.4605
7	'felicidad'	0.4598
8	'institucionales'	0.4591
9	'esencia'	0.4565
10	'celebración'	0.45

españolas como las originarias de América, el portugués o el inglés [19]. The mention of indigenous languages of the Americas and other international languages (Portuguese, English) enables the application of the strategy of global coexistence and the tactic of universalisation, which emphasises combining efforts to develop intercultural dialogue.

During the analysis, the key token identified was *hispanico*, a generalising word encompassing all aspects related to Spain and Spanish-speaking countries, thus reinforcing the context of harmonisation: *El Rey abre el encuentro de Cádiz, punto de encuentro de la cultura hispanica y dedicado a estudiar esta vez el idioma como fenómeno de mestizaje* [19]. The concept *mestizaje*, which describes the combination of different cultures and the subsequent emergence of a new one, merits particular attention, and in the following sentence it foregrounds the themes of continuity and collective memory: *Varios siglos después, «el Diccionario de Oxford reconoce el origen español de la palabra mestizaje»* [19]. The other tokens that proved to be fairly frequent – 'naciones', 'seguridad', 'identidad' – similarly to those identified in the French-language examples, support the strategies of global coexistence, legitimisation, and identification: the author employs evaluative markers that indicate an intention to accentuate the importance of the event and the continuation of intercultural co-operation.

At the same time, the precedent-related cultural phenomena and allusions uncovered in the article serve to connect the past with the present, evoking the context of linguistic creativity traditionally associated with Cádiz (*Zacatecas, Valladolid, Rosario, Cartagena de Indias, Valparaíso, Panamá, San Juan de Puerto Rico, Córdoba (Argentina)*) – all of these being the cities that served as venues for previous International Congresses of the Spanish Language; *Cádiz C.F.* as the name of the football club), an expression (*si me queréis, irse*), and a situation (*la época del Inca Garcilaso de la Vega*).

In this connection, to examine the actualisation of the dominant linguocultural concept LENGUA within the Mexican media discourse, the article *La Academia Mexicana de la Lengua y la UNAM* dated 15 September 2025 (*La Jornada, Mexico*) [20] was subjected to content analysis. It was written on the occasion of the 150th anniversary of the Mexican Academy of Language, highlighting the collaboration of this institution with the National Autonomous University of Mexico. The most frequent tokens identified after text processing are presented in Figure 4.

To deepen our understanding of these results, we also conducted an analysis of the key concepts derived from the article, among which the most frequent include 'relacionarse', 'mexicanismo',

'pertenencia', 'pluralidad', 'nación', 'identidad', 'institución', 'unidad', 'mexicano', 'lingüístico' (see Table 4).

In like manner to the foregoing example, the conceptual exponent *lengua* proved to be among the most frequent lexemes: *Nuestra lengua, el español, es más que un vehículo de entendimiento: refleja lo que somos, nos sitúa, orienta y ancla en el mundo; es la condición y la herramienta más humana porque nos permite comprendernos, relacionarnos y dar sentido a nuestros actos* [20]. It is worth observing that in the cited passage the manifestation of this concept is bound up with the realisation of several strategies, namely, those of cultural legitimisation, identification, and global coexistence. This connection is achieved through the author's recourse to first-person plural forms (*nosotros*), as well as the use of the verbs *comprendernos* and *relacionarnos*, which reinforces the idea of unity. As within the Spanish media discourse, one may find cases of introducing the lexeme *idioma*: *Concebida como una institución que resguarda, estudia y fomenta el idioma que hoy nos hermana a más de 600 millones de hablantes, ha sostenido un empeño constante, un diálogo creador y un servicio insustituible a México y a la comunidad hispanica global* [20]. In this sample, the presence of numerical indicators serves as quantitative corroboration, but they also highlight the international status of the language and convey the idea of a dialogue between Mexico and the entire Spanish-speaking community.

Another nationally marked concept that was identified during the analysis is the token 'mexicanismo', which ensures the strategy of identification and, as such, is associated with linguistic and cultural identity: *Sin duda, México ha aportado al español su riqueza léxica e imaginarios, plasmada en obras como el Diccionario breve de mexicanismos, entre otros, que reconocen y estimulan la creatividad de la sociedad mexicana y la legitimidad de sus expresiones más propias y genuinas* [20]. At the level of linguocultural concepts, the strategies of identification and continuity are actualised by the lexemes *pertenencia, patrimonio, historia, legado, and identidad*: *Encarnaron una doble pertenencia y supieron tender puentes; El lazo entre ambas instituciones es consustancial a la trama intelectual, literaria e histórica del país, ya que abrazan una misma convicción: la lengua como patrimonio y como germen de la libertad* [20]. Other lexemes – *nación, unidad, institución*, as well as *sociedad* – are correlated with the above-mentioned strategies of institutionalisation, cultural continuity, and harmonisation, serving to strengthen the argument aimed at supporting the main idea of the article, which is celebrating the many years of co-operation between both the two institutions and representatives of the Spanish-speaking world in general.

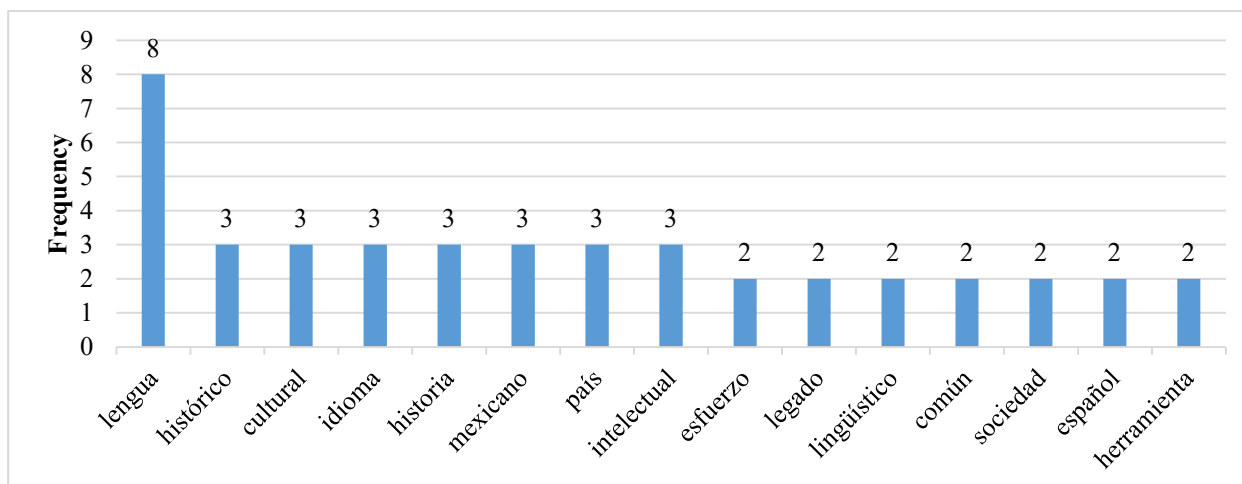
Fig. 4. 15 most frequent tokens in the text (the article from *La Jornada*)

Table 4

Selected dominant concepts identified in the article from *La Jornada*

Nº	Key concept	Relevance index
1	'relacionamos'	0.5572
2	'mexicanismos'	0.5346
3	'pertenencia'	0.524
4	'pluralidad'	0.523
5	'naciones'	0.5186
6	'identidad'	0.5116
7	'instituciones'	0.5109
8	'unidad'	0.5107
9	'mexicanos'	0.5066
10	'mexicana'	0.4954

In the final analysis, the examination of the articles gives rise to the following considerations. (1) The DLC LANGUE / LENGUA reveal a multifaceted nature in the establishment of associative links with other concepts that stand to them in closer or remoter lexical-semantic relations. (2) Most of the identified tokens and dominant concepts, which in this study are treated as conceptual exponents, can be combined into macroconcepts NATION / NACIÓN, IDENTITÉ / IDENTIDAD, CULTURE / CULTURA, COOPERATION / COOPERACIÓN, INSTITUTION / INSTITUCIÓN, which are integrated into both the lexical-semantic and cognitive space of the DLC LANGUE / LENGUA. (3) Some of the identified concepts and communicative strategies are realised across the texts from the periodicals of all countries, whereas certain others appear solely in a text from a single periodical. (4) The lexical exponents of concepts common to the French-language media discourse prove to be *français, francophonie, monde, québécois*, while those occurring within the Spanish-language discourse are *español, cultura, historia, nación, relación, institución, identidad*. (5) At the level of cross-lingual comparison, the shared concepts may be grouped by the following pairs: *institution – institución, linguistique – lingüístico, culture – cultura, nation – nación, moyen – herramienta*, alongside the provisional pairs *diversité – pluralidad, sauvegarde – seguridad* (cf. the verb *salvaguardar* attested in one of the texts). (6) At the level of discourse analysis, the strategies common to all the texts examined are those of cultural legitimisation, cultural identification, cultural continuity, global coexistence, and institutionalisation. On the contrary, the strategies specific

to the discourses of individual countries are those of solidarity and protection (Canada), demographic forecasting (France, Spain), and harmonisation (Spain, Mexico). This allows us to conclude that the French-language and Spanish-language media discourses convey a common idea of the key role of language in national self-identification and establishing international ties. At the same time, what is distinctive of the Canadian periodical (and, to some extent, of the French one) is the attention given to issues of protecting the French language and cultural heritage (cf. the lexemes *sauvegarde, culture, avenir*), as well as to the current challenges arising in this context. In the Spanish-language periodicals, by contrast, such concerns have not been found to be expressed directly; instead, they contain reflections on the further development of linguistic co-operation. Yet in the Mexican discourse the theme of national self-identification is likewise foregrounded, accompanied by the delineation of specific cultural markers (*mexicanismo*), which, in turn, aligns it with the Canadian text (*québécois*). (7) A number of identified precedent-related phenomena and allusions are nationally marked and function to call upon cultural facts, not merely appealing to the reader's erudition, but also strengthening the author's arguments and strategies employed.

That notwithstanding, it must be acknowledged that the analytical framework applied requires further corroboration on a more extensive textual corpus, so as to allow for more generalised conclusions as regards the realisation of particular DLC within media discourse, as well as to identify tendencies in the actualisation of specific concepts while bringing to light a greater number thereof.

Conclusions and prospects. With everything considered, the comparative content analysis proved to reveal the significant role that the DLC LANGUE / LENGUA play in the conceptual organisation and actualisation of cultural experience in the texts drawn from the selected media. The study confirmed the assumption of their generally identical realisation in both languages, whilst enabling the determination of diverse specific features that they acquire in the periodicals of each country.

Further avenues of research present themselves in the continued testing of content analysis on a larger corpus with the involvement of additional Python libraries and various linguistic methods. This might be profitably carried out to reinforce the validity of the results derived from the integration of quantitative and qualitative approaches to examining the linguocultural parameters of media discourse.

Bibliography:

1. Van Dijk T. A. News as discourse. Hillsdale (NJ) : Lawrence Erlbaum Associates, 1988. 208 p.
2. Fairclough N. Media discourse. London : Arnold, 1995. 214 p.
3. Fowler R. Language in the News: Discourse and Ideology in the Press. London : Routledge, 1991. 254 p.
4. Palmer G. B. Toward a Theory of Cultural Linguistics. Austin : University of Texas Press, 1996. 360 p.
5. Sharifian F. Cultural Conceptualisations and Language. Amsterdam ; Philadelphia : John Benjamins, 2011. 238 p.
6. Montgomery M. Language, Media and Culture: The Key Concepts. London : Routledge, 2019. 180 p.
7. Burger M. L'analyse linguistique des discours médiatiques. Québec : Nota Bene, 2008. 290 p.
8. Taylor C., Del Fante D. Comparing across languages in corpus and discourse analysis: some issues and approaches. *Meta*. 2020. Vol. 65, No. 1. Pp. 29–50. DOI: <https://doi.org/10.7202/1073635ar>
9. Nguyen D. et al. How We Do Things With Words: Analyzing Text as Social and Cultural Data. *Frontiers in Artificial Intelligence*. 2020. Vol. 3, article 62. DOI: <https://doi.org/10.3389/frai.2020.00062>
10. Gallardo Pauls B. La necesidad de modelos en lingüística: sobre el análisis del discurso digital. *Pragmalingüística*. 2023. Núm. 31. Pp. 127–151. DOI: <https://doi.org/10.25267/Pragmalinguistica.2023.i31.06>.
11. Загнітко А., Богданова І. Лінгвокультурологія : навч. посіб. для студ. закл. вищ. осв. / за ред. А. Загнітка. 4-те вид., перероб. і доп. Вінниця : ТВОРИ, 2024. 264 с.
12. Сасевич І. Г. Комплексна методика виокремлення лінгвокультурних домінант. *Studia Philologica*. 2014. Вип. 1. С. 95–99. URL: <https://studiap.kubg.edu.ua/index.php/journal/article/view/60>
13. Яницька О. М. Лінгвокультурологічні аспекти аналізу концепту ВІРА в англійській, українській та французькій мовах. *Науковий вісник Херсонського державного університету. Серія: Перекладознавство та міжнародна комунікація*. 2016. Вип. 6. С. 117–121. URL: <https://tsj.journal.kspu.edu/index.php/tsj/article/view/413>
14. Corpus Linguistics, Context and Culture / Ed. by V. Wiegand, M. Mahlberg. Berlin : De Gruyter, 2020. 492 p. DOI: <https://doi.org/10.1515/9783110489071>
15. The Routledge Handbook of Corpus Approaches to Discourse Analysis / Ed. by E. Friginal, J. A. Hardy. London : Routledge, 2020. 662 p. DOI: <https://doi.org/10.4324/9780429259982>
16. Keller D. Programming for Corpus Linguistics with Python and Dataframes. Cambridge : Cambridge University Press, 2024. 114 p.
17. Lemoigne V. À quoi ressemblera la langue française de demain ? *Le Figaro*. Le 21 mars 2024. URL : <https://www.lefigaro.fr/langue-francaise/actu-des-mots/a-quoi-ressemblera-la-langue-francaise-de-demain-20240321>
18. Biron M. et al. Célébrons notre langue française, celle qui nous unit et nous définit. *Le Devoir*. Le 18 mars 2023. URL : <https://www.ledevoir.com/opinion/libre-opinion/785803/libre-opinion-celebrons-notre-langue-francaise-celle-qui-nous-unit-et-nous-definit>
19. Alemany L. Don Felipe inaugura el Congreso de la Lengua: «El siglo XXI debe ser el siglo del español». *El Mundo*. 27 de marzo de 2023. URL: <https://www.elmundo.es/cultura/literatura/2023/03/27/64218071fc6c832a308b45a4.html>
20. Lomeli Vanegas L. La Academia Mexicana de la Lengua y la UNAM. *La Jornada*. 15 de septiembre de 2025. URL: <https://www.jornada.com.mx/noticia/2025/09/15/cultura/la-academia-mexicana-de-la-lengua-y-la-unam>

Баранцев Я. Порівняльний контент-аналіз реалізації лінгвокультурних домінант LANGUE / LENGUA у французькомовному іспанськомовному медіадискурсах (на матеріалі періодичних видань Франції, Канади, Іспанії та Мексики)

Анотація. У статті здійснено спробу огляду й апробації порівняльного контент-аналізу реалізації лінгвокультурних домінант LANGUE / LENGUA у французькомовному й іспанськомовному медіадискурсах на матеріалі фрагментів, відібраних зі статей періодичних видань «Le Figaro» (Франція), «Le Devoir» (Канада), «El Mundo» (Іспанія) і «La Jornada» (Мексика). Закцентовано, що концепт МОВА належить до фундаментальних у національній картині світу кожного народу, з огляду на що постає питання про зіставлення сутнісних ознак його актуалізації в мовах світу, зокрема таких, як французька й іспанська, що їх об'єднує не лише міжнародний статус, але й спільність типологічних рис на всіх рівнях мовної ієрархії. Відповідно, у розвідці здійснено перевірку припущення, яке ґрунтується на загальній подібності реалізації лінгвокультурних домінант LANGUE / LENGUA у французькомовному й іспанськомовному медіадискурсах на прикладі періодичних видань декількох країн на базі порівняльного контент-аналізу.

Конкретніше, для аналізу був використаний комплекс бібліотек мови програмування Python, що забезпечив цикл попередньої обробки тексту, з-поміж яких – NLTK, StopwordsISO, Stanza, KeyBERT, Pandas і Matplotlib. Окрім того, дані, отримані внаслідок використання зазначених комп'ютеризованих інструментів, були доповнені результатами залучення методів семантичних полів (для об'єднання слів за тематичними групами-макронцептами) і дискурс-аналізу, що дало змогу виявити й описати імпліцитні стратегії, які реалізуються через повторення окремих культурно маркованих ключових слів. У процесі дослідження з'ясувалося, що домінанти LANGUE / LENGUA виконують вагомую роль у концептуальній організації, актуалізації культурного досвіду й формулюванні аргументації в текстах відібраних періодичних видань.

У зв'язку з цим на основі аналізу текстів на рівні концептуальної мережі та дискурсивних стратегій зроблено висновок про те, що у французькомовному й іспанськомовному медіадискурсах реалізується спільна ідея ключової ролі мови в національній самоідентифікації та встановленні міжнародних зв'язків. Специфічним для текстів видань Канади та Франції виявилось порушення проблем захисту французької мови, збереження культурної спадщини, тоді як в іспанськомовних виданнях актуалізується ідея подальшого розвитку мовної співпраці, хоча в мексиканському медіадискурсі також зафіксовано тему національної самоідентифікації з визначенням питомих маркерів,

що, своєю чергою, об'єднує його з канадським. Рівночасно з тим установлено, що культурно марковані прецедентні феномени підсилюють авторську аргументацію та ефект залучених стратегій. Зрештою, окреслено перспективи подальшої апробації аналізу та наголошено на необхідності його застосування на ширшому текстовому матеріалі із залученням додаткових бібліотек Python і лінгвістичних методів для підвищення репрезентативності результатів дослідження.

Ключові слова: медіалінгвістика, лінгвокультурна домінанта, контент-аналіз, мовна картина світу, романські мови.

Дата першого надходження рукопису
до видання: 10.11.2025

Дата прийнятого до друку рукопису
після рецензування: 08.12.2025

Дата публікації: 31.12.2025