

DISCOURSE AND ITS ROLE IN LINGUISTICS TYPES AND CLASSIFICATIONS

<https://doi.org/10.5281/zenodo.23254028>

Maftuna Inomjanova Abduxamid qizi

International School of Finance and Technology (ISFT), English Language Teacher

inomjanovamaftuna64@gmail.com

Abstract

This article examines discourse as situated language use and evaluates the criteria employed to classify it. It asks whether medium, institutional setting and communicative purpose can be treated as competing types within a single taxonomy. A conceptual comparison of linguistic approaches is combined with a transparent analysis of twelve constructed discourse cases. Each case is coded independently for medium, interaction, domain and dominant purpose, while published corpus documentation supplies a separate empirical reference point. The analysis shows how a classroom exchange, public announcement or online explanation can receive several compatible labels without becoming a contradictory case. The methodological contribution is a layered classification procedure that separates the object being described from the analytical perspective applied to it. The article recommends explicit coding rules, attention to mixed cases and cautious interpretation of numerical summaries drawn from illustrative material.

Keywords

discourse; classification; register; genre; interaction; corpus design; linguistic analysis

Annotatsiya

Ushbu maqolada diskurs vaziyatga bog'liq til qo'llanishi sifatida o'rganilib, uni tasniflash mezonlari tahlil qilinadi. Tadqiqot kommunikatsiya vositasi, institutsional muhit va muloqot maqsadini bitta tasnifdagi o'zaro raqobatlashuvchi turlar sifatida talqin qilish mumkinligini tekshiradi. Nazariy yondashuvlar qiyosiy tahlili maqola uchun tuzilgan o'n ikkita diskurs misolining ochiq tahlili bilan birlashtiriladi. Har bir misol vosita, o'zaro ta'sir, soha va asosiy maqsad bo'yicha alohida kodlanadi, e'lon qilingan korpus ma'lumotlari esa mustaqil empirik manba vazifasini bajaradi. Tahlil bir misolga bir nechta mos tasnif belgisi berilishi mumkinligini ko'rsatadi. Metodologik hissa tavsiflanayotgan obyektga unga

qo'llanayotgan tahliliy yondashuvdan ajratadigan ko'p qatlamli tasnif tartibidan iborat. Maqolada aniq kodlash qoidalari, aralash holatlarga e'tibor va namunaviy material statistikalarni ehtiyotkor talqin qilish tavsiya etiladi.

Kalit so'zlar

diskurs; tasnif; registr; janr; o'zaro ta'sir; korpus tuzilishi; lingvistik tahlil

Аннотация

В статье дискурс рассматривается как использование языка в конкретной ситуации и анализируются критерии его классификации. Исследуется, можно ли считать канал коммуникации, институциональную среду и коммуникативную цель конкурирующими типами одной классификации. Сопоставление теоретических подходов дополняется прозрачным анализом двенадцати специально составленных примеров дискурса. Каждый пример отдельно кодируется по каналу, взаимодействию, сфере и основной цели, а опубликованная документация корпусов служит самостоятельным эмпирическим ориентиром. Анализ показывает, что одному случаю могут соответствовать несколько совместимых классификационных обозначений. Методологический вклад состоит в многоуровневой процедуре, отделяющей описываемый объект от применяемого аналитического подхода. Рекомендуются явные правила кодирования, внимание к смешанным случаям и осторожная интерпретация статистических сводок иллюстративного материала.

Ключевые слова

дискурс; классификация; регистр; жанр; взаимодействие; структура корпуса; лингвистический анализ

Introduction

A university lecture is spoken, educational, institutionally organised and usually explanatory. A lecture recorded for an online platform also becomes a digital artefact that viewers can pause, annotate and discuss. None of these descriptions cancels the others. The difficulty begins when a classification presents "spoken discourse", "educational discourse" and "digital discourse" as if they were mutually exclusive alternatives. Such a list confuses the channel carrying an event with the social setting and the purpose of the event.

The object of this study is discourse understood as language used within a communicative situation. Its subject is the relationship between the criteria used to describe discourse types. The aim is to develop a practical classification procedure

that preserves differences among medium, interaction, domain and purpose. Three tasks follow: compare influential linguistic perspectives, apply an explicit coding scheme to contrasting cases, and assess what descriptive counts can and cannot establish.

Discourse matters to linguistics because a grammatical description alone does not explain how a contribution becomes relevant, how a participant obtains a turn, or why the same wording can sound appropriate in one setting and intrusive in another. Brown and Yule place language use at the centre of discourse analysis [1]. The present argument proceeds from that orientation but concentrates on a narrower methodological problem: the classification of a communicative event should remain inspectable when several labels are simultaneously valid.

The proposed contribution is a synthesis for analytical and teaching purposes, rather than a claim to have discovered previously unknown discourse categories. It is especially useful in small research projects where investigators must specify what their sample represents before comparing linguistic features. A classification that conceals its own criteria can produce numerical findings that look precise while describing an uncertain object.

Literature review

Different traditions select different aspects of discourse. Text-oriented work examines reference, lexical continuity and relations between stretches of language. Halliday and Hasan provide a foundational account of cohesion [2]. Interactional work attends to the organisation of participation, while ethnographic approaches ask how a communicative activity fits a community's expectations [5]. Critical analysis connects textual choices with institutional arrangements and social consequences [6]. These traditions offer analytical lenses; they do not necessarily identify separate populations of discourse.

Register and genre also need to be distinguished. Register analysis relates recurring linguistic patterns to situational conditions [4]. Genre analysis considers conventional communicative activities and their organisation [7]. A research article and a job application can both be written and formal while belonging to different genres. Conversely, a single educational domain contains seminars, lectures, examination instructions and informal consultations. The distinction prevents a broad domain label from doing the work of a more specific description.

Biber's study examined twenty-three spoken and written genres and identified six dimensions of variation [3]. This evidence supports a multidimensional view of linguistic differentiation; it does not imply that any locally devised set of six categories reproduces those statistical dimensions. The present framework

therefore makes no factor-analytic claim. Its four coding axes are practical descriptors chosen for the constructed examples.

Research resources add another perspective. The RST Discourse Treebank contains 385 Wall Street Journal articles annotated for rhetorical structure [9]. Its scope is valuable for studying relations within journalistic texts, but a news sample cannot automatically stand for conversational or classroom interaction. An adequate classification must consequently describe both the phenomenon and the restrictions imposed by the available material.

Research methodology

The study combines a focused conceptual review with a worked classification exercise. The literature was selected purposively to represent textual, register, genre, interactional and social approaches; this is not a systematic review or a bibliometric census. Twelve short cases were constructed for this article. They are analytical examples, not extracts from interviews, classrooms or an observed population. No participants were recruited, no corpus was newly collected, and no inter-annotator reliability estimate is claimed.

Each case receives one primary label on four axes. Medium records how the contribution is presented: spoken, written or multimodal. Interaction distinguishes contributions organised around responsive turns from contributions designed primarily for extended delivery. Domain specifies the relevant activity setting. Purpose identifies the dominant action in the stated scenario: informing, persuading, instructing or maintaining a relationship. A primary label is a coding convenience, not a denial of secondary functions.

Counts are computed directly from the cases printed in Table 2. A category share is $p = 100 \times n / 12$, where n is the number of cases assigned that label. These percentages describe the composition of the constructed set only. They cannot estimate how often a discourse type occurs in society. Published corpus design information is reported separately and is not merged into the example counts.

Table 1. Independent axes for describing discourse

Axis	Question	Primary coding rule
Medium	How is meaning presented?	Spoken, written or multimodal in the stated event
Interaction	How can participants respond?	Responsive turns versus extended delivery
Domain	What activity is taking place?	Educational, public, academic, commercial or interpersonal

Axis	Question	Primary coding rule
Purpose	What is the main action?	Inform, persuade, instruct or sustain a relationship

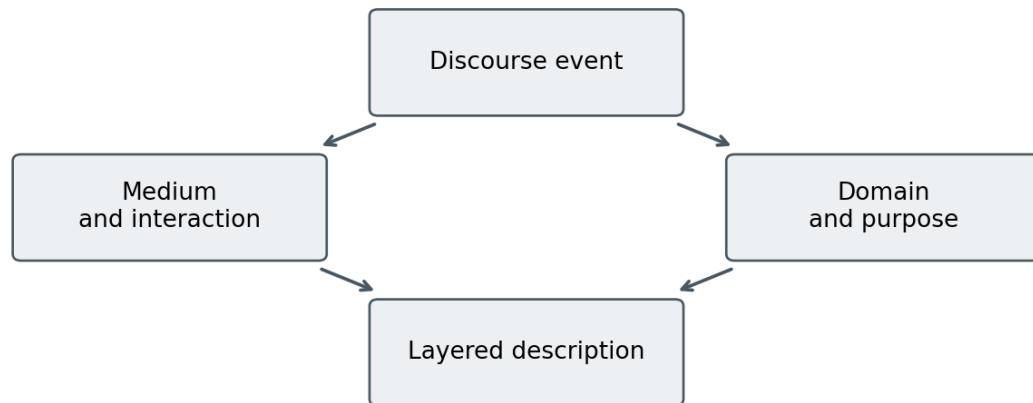


Figure 1. A layered description of a discourse event. Conceptual diagram developed for this article.

Analysis and results

A transparent set of classification cases

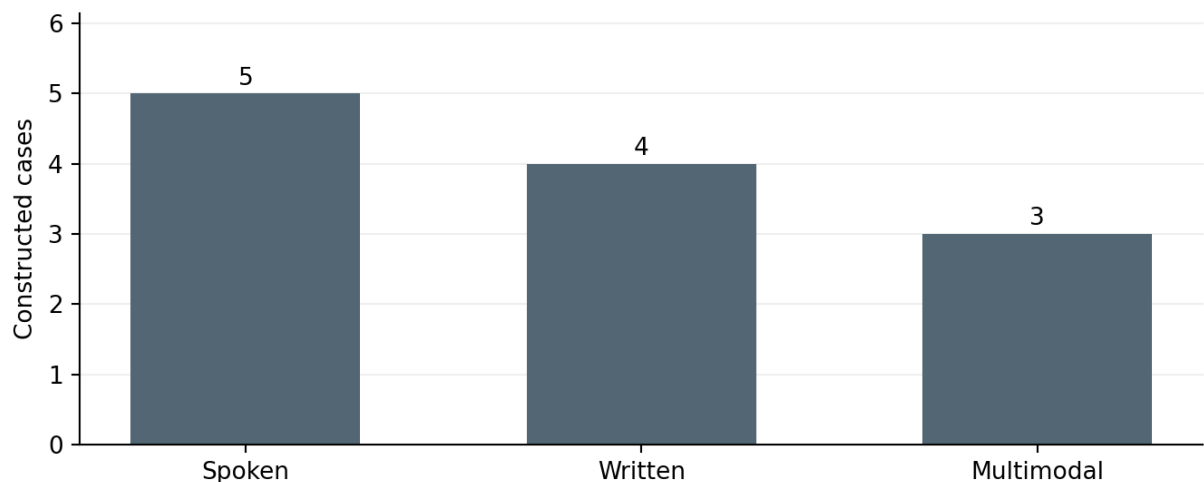
Table 2. Complete constructed case set and primary codes

ID	Scenario	Medium	Interaction	Purpose
D01	Tutor answers a question	Spoken	Dialogic	Inform
D02	Friends exchange greetings	Spoken	Dialogic	Relate
D03	Candidate debates a policy	Spoken	Dialogic	Persuade
D04	Teacher explains a procedure	Spoken	Monologic	Instruct
D05	Station announces a delay	Spoken	Monologic	Inform
D06	Article explains a finding	Written	Monologic	Inform
D07	Manual lists safety steps	Written	Monologic	Instruct
D08	Editorial supports a policy	Written	Monologic	Persuade
D09	Friends exchange text greetings	Written	Dialogic	Relate
D10	Video shows an experiment	Multimodal	Monologic	Inform
D11	Advert combines text and images	Multimodal	Monologic	Persuade
D12	Live screen-sharing tutorial	Multimodal	Dialogic	Instruct

The cases clarify why “spoken” is insufficient as a functional description. D02 maintains a social relationship, D03 advances a position, and D04 guides an action. All are spoken, but their participation arrangements and communicative goals differ. The reverse comparison is equally instructive: persuasion occurs in a spoken debate, a written editorial and a multimodal advertisement. A classification organised around only one of these properties would conceal the other contrasts.

The written exchange in D09 is dialogic because one participant’s greeting creates an opportunity for another participant to respond. Writing does not imply the absence of interaction. D12 is multimodal because the explanation relies on a shared screen as well as speech; it remains dialogic because the learner can ask questions during the demonstration. The classification records these properties without treating digital mediation as an independent communicative purpose.

The domain codes, omitted from Table 2 to keep its columns readable, are educational for D01, D04, D07, D10 and D12; public for D03, D05 and D08; interpersonal for D02 and D09; academic for D06; and commercial for D11. Thus every code used in the summary remains recoverable from the printed scenarios and the explicit domain assignment.



Source: Table 2; n = 12; illustrative composition, not population prevalence.

Figure 2. Primary medium codes in the constructed case set.

Table 3. Descriptive summary of the constructed cases

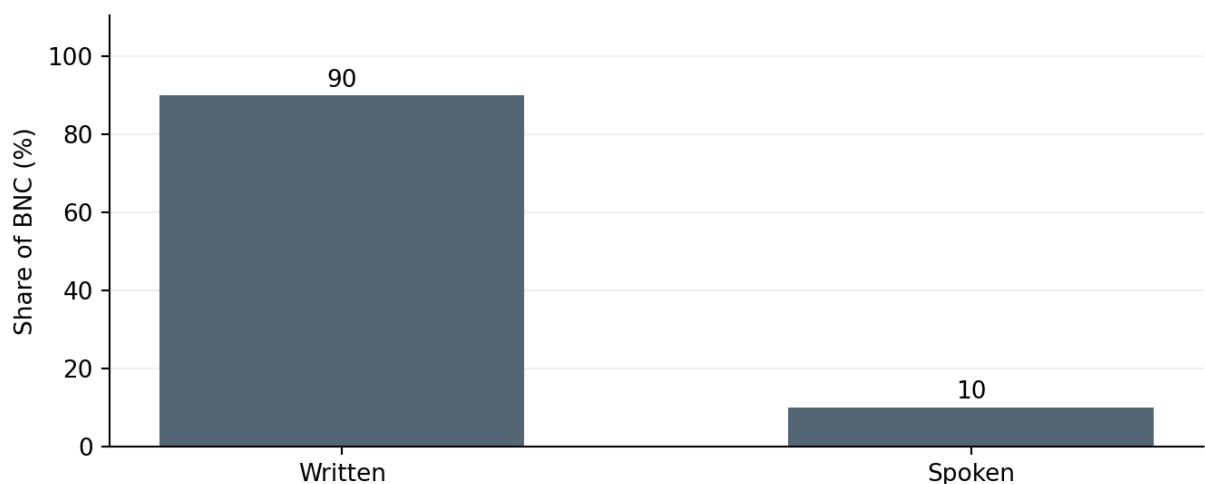
Axis	Category	Count	Share
Medium	Spoken	5	41.7%
Medium	Written	4	33.3%
Medium	Multimodal	3	25.0%
Purpose	Inform	4	33.3%

Axis	Category	Count	Share
Purpose	Relate	2	16.7%
Purpose	Persuade	3	25.0%
Purpose	Instruct	3	25.0%

Five cases are spoken, four written and three multimodal, corresponding to 41.7%, 33.3% and 25.0% of the set. The counts sum to twelve because a single primary medium was assigned to each case. Purpose is distributed differently: informing accounts for four cases, persuasion and instruction for three each, and relationship maintenance for two. These totals demonstrate that the coding axes partition the same set in different ways. They do not test the empirical importance of any category.

A richer coding scheme could allow several purposes per event. An advertisement may inform while persuading; a teacher's instruction may also encourage a hesitant learner. In that design, category totals could exceed the number of events. Reporting a percentage without stating whether labels are exclusive would then be misleading. This is a concrete reason to document the coding rule before presenting a diagram.

Published corpus composition as a sampling constraint



Source: University of Oxford BNC documentation [10]; published corpus composition.

Figure 3. Written and spoken shares of the British National Corpus [10].

The British National Corpus documentation reports a 90% written and 10% spoken composition [10]. These published proportions concern that resource, whereas Figure 2 describes constructed cases. Neither distribution should be read as the proportion of writing or speech in everyday human communication. Their

contrast instead illustrates how sampling decisions shape the evidence available to an analyst.

Suppose an investigator studies second-person pronouns in a collection dominated by written academic material. A low overall rate could reflect sample composition rather than the linguistic insignificance of direct address. Comparisons should therefore retain domain and interaction metadata and use comparable denominators. A count per text answers a different question from a count per thousand words; both differ from a count per conversational turn. Classification is part of measurement design, not merely an introductory vocabulary exercise.

Discussion

The layered procedure improves conceptual clarity but does not solve every boundary problem. “Academic” and “educational” may overlap when a researcher teaches specialist material. A recorded seminar may contain both extended delivery and responsive discussion. A social media post may become one contribution to a larger exchange even when the post itself contains no visible response. The unit of analysis must therefore be declared: message, segment, whole event or recurring practice.

Analytical approaches should also remain distinguishable from object labels. A critical discourse analysis can investigate an educational exchange; a conversation analysis can investigate a commercial service encounter. Treating “critical discourse” as a category parallel to “commercial discourse” would confuse the investigator’s perspective with the participants’ activity. The present classification keeps that distinction explicit while recognising that social power may affect all domains [6].

For classroom use, students can classify the same event independently on each axis and justify the evidence for their choices. Differences in coding become productive when they concern identifiable features rather than an undefined intuition about formality. A useful follow-up is to change one property of a case: replace a live tutorial with a recorded video or a station announcement with a text notification. Students can then identify which descriptors change and which remain defensible.

For research, the next step is validation on authentic data. Investigators would need a sampling frame, a coding manual, independent annotation and a procedure for resolving disagreements. The illustrative counts here provide none of those guarantees. They demonstrate a procedure that can be tested, rather than a population model whose accuracy has already been established.

Conclusion and recommendations

Discourse classification is most informative when it separates the properties of an event into explicit descriptive axes. Medium, interaction, domain and purpose can coexist in one account without being reduced to competing types. The worked cases reveal predictable overlaps and make the consequences of primary-label coding visible.

Researchers should define their analytical unit, state whether categories are exclusive, publish the evidence supporting assignments, and distinguish sample composition from claims about language use generally. Teachers can use the same procedure to connect linguistic forms with communicative activities. Future empirical work should test the scheme on authentic spoken, written and multimodal material and report disagreement rather than hiding difficult cases. A classification is useful to the extent that another analyst can understand how it was applied.

Interaction can additionally be examined through turn-taking organisation [11], while multimodal cases require attention to visual meaning [12]. Rhetorical relations offer a complementary account of textual organisation [8].

REFERENCES:

- [1] Brown, G., & Yule, G. (1983). *Discourse Analysis*. Cambridge University Press.
- [2] Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. Longman.
- [3] Biber, D. (1988). *Variation across Speech and Writing*. Cambridge University Press. <https://www.cambridge.org/core/books/variation-across-speech-and-writing/A546CF5ED8F8E62F1432CB2F369CF356>
- [4] Biber, D., & Conrad, S. (2009). *Register, Genre, and Style*. Cambridge University Press.
- [5] Hymes, D. (1974). *Foundations in Sociolinguistics: An Ethnographic Approach*. University of Pennsylvania Press.
- [6] Fairclough, N. (1992). *Discourse and Social Change*. Polity Press.
- [7] Swales, J. M. (1990). *Genre Analysis: English in Academic and Research Settings*. Cambridge University Press.
- [8] Mann, W. C., & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text*, 8(3), 243–281. <https://doi.org/10.1515/text.1.1988.8.3.243>

[9] Carlson, L., Marcu, D., & Okurowski, M. E. (2002). RST Discourse Treebank. Linguistic Data Consortium, LDC2002T07. <https://catalog.ldc.upenn.edu/LDC2002T07>

[10] University of Oxford. (n.d.). About the British National Corpus. <https://www.natcorp.ox.ac.uk/corpus/index.xml> (accessed 7 October 2026).

[11] Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50(4), 696–735. <https://doi.org/10.2307/412243>

[12] Kress, G., & van Leeuwen, T. (2006). *Reading Images: The Grammar of Visual Design* (2nd ed.). Routledge.