

Contrastive Discourse Analysis

Niall Curry, Manchester Metropolitan University, Manchester, United Kingdom

© 2026 Elsevier Ltd. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

Introduction	1
Contrastive Linguistics	1
Key Concepts and Data in Contrastive Studies	2
Case Study: Corpus-Based Contrastive Discourse Analyses of Public-Oriented Research Communication	3
Conclusion	5
References	5

Key Points

- This article traces the evolution of contrastive linguistics over time and presents current work in corpus-based contrastive linguistics.
- It discusses core theoretical and methodological principles, including comparability, the *tertium comparationis*, and equivalence testing, demonstrating how these concepts enable principled, falsifiable cross-linguistic comparison.
- It showcases the affordances of corpus-based contrastive discourse analysis through a case study examining how global crises are discursively constructed across languages.
- It reflects on the potential for corpus-based contrastive discourse analysis to advance contemporary work in discourse studies.

Abstract

In this article, I review the field of contrastive linguistics, tracking its development over time, its theoretical underpinnings, and its integration with corpus linguistics. I then offer a reflective case study, designed to illustrate how contrastive approaches can shed light on the varied discursive construction of shared societal challenges across languages. Specifically, through a focus on public-oriented research communication in English, French, and Spanish, I demonstrate how language, culture, and situated epistemologies shape knowledge-making in different linguistic and cultural contexts. In closing, I point to future directions in contrastive discourse studies.

Introduction

In this article, I discuss the field of contrastive discourse studies, reflecting on its origins, its development over time, and the key theoretical and methodological concepts that have underpinned this development. In so doing, I draw particular attention to theoretical concepts in contrastive linguistics and the synergetic affordances of corpus linguistics thereof, from a methodological perspective. This discussion is followed by a reflective case study, designed to operationalize these notions through an interrogation of the use of contrastive approaches in corpus-based discourse studies. In concluding, I consider the affordances of such contrastive work and I point to future directions for corpus-based contrastive discourse studies.

Contrastive Linguistics

Contrastive linguistics is a branch of linguistics research concerned with the systematic comparison of two or more languages. Research in this area typically seeks to identify and explain patterns of similarity and difference in terms of facets of language including form, meaning, and use. Contemporary contrastive work is socially situated, guided by a volition to bridge communicative cultures, inform translation activities, reshape language education, and unpack culturally and socially situated perspectives on shared issues. Regardless of the area of application, conducting such contrastive work requires a high degree of theoretical and methodological sophistication, as any findings that tell us something about how different languages correspond to one another must be established through principled comparison. As such, in its current conception, contrastive linguistics is underpinned by rigorous, empirical foundations that draw from both contrastive and corpus linguistics. Yet, this position was hard earned as, as a field, the origins of contrastive linguistics lie in critiques of methodological circularity and unfounded pedagogical applications.

Though contrastive work can be dated back to now ancient civilizations, it was in the mid-twentieth-century that the contrastive analysis approach was formalized through the seminal works of Fries (1945) and Lado (1957). Both Fries and Lado shared a view that contrastive analysis could change the way languages were taught and learned. This argument was premised on the claim that the systematic comparison of a learner's first and target languages could inform language teaching, predict errors, and anticipate areas of learning difficulty. Operationally, the contrastive analysis hypothesis appeared logical, being based on the assumption that those features of a language that are similar in two languages (e.g., cognate words in English and French) could be leveraged to facilitate language learning and that those areas in which languages differ (e.g., continuous verb forms in English and French) would cause interference during the learning process due to a lack of formal correspondence. This hypothesis had mass appeal, emerging at a time of increased need for language competences (following two world wars and growing global and diplomatic exchange) and amid a rising preoccupation with educational theory in the academy. However, this early phase in the journey of contrastive studies proved methodologically and theoretically fragile. The capacity for the approach to accurately predict errors was inconsistent, and although empirical data was sometimes used in contrastive analyses, engagement with language data typically ignored context and drew on poor research design. As a result, contrastive analysis fell from favor (Corder, 1981; Klein, 1986), only to be reborn in the 1980s and invigorated in the 1990s, following crucial theoretical and methodological developments.

From the 1980s onwards, the field underwent a substantial reconceptualization. Rather than attempting to predict learner errors, contrastive linguistics re-emerged as a descriptive and explanatory discipline grounded in authentic language use, functional interpretation, and corpus evidence (Altenberg & Granger, 2002; James, 1980). This renewal was built on critical concepts of comparison, comparability, and sameness, and aligned contrastive linguistics closely with corpus linguistics, whose tools and methods enabled large-scale, replicable analyses of naturally occurring comparable language data. Johansson (2012) characterizes this shift as "contrastive linguistics in a new key" (p. 46), constituted by a move away from narrow teaching applications toward text-based comparison, theoretically explicit notions of comparability, and the systematic exploitation of multilingual corpora. As such, contemporary contrastive linguistics is less concerned with predicting language learning difficulties and is more concerned with understanding how languages (often with a focus on specialized domains of use) correspond in terms of comparable communicative contexts.

The defining feature of modern contrastive studies is its commitment to principled comparison. Earlier work on contrastive analysis was plagued by issues of methodological rigor, as comparisons were often designed in ways that did not engage with processes of falsification. This meant that if a researcher were to compare two languages and focus on something they know to be shared across them (e.g., forms of cognate words), they could conclude that this lexical item corresponds across the languages studied. Yet, we know that a word is more than its form. For example, if we consider how Latinate cognate forms in English are often used in more formal registers than they are in languages like French, Italian, and Spanish, we may note that while a form is shared, the contexts in which a word is used can differ greatly. This kind of thinking revolutionized approaches to contrastive studies as, in the determination of sameness or lack thereof, the field became preoccupied with two things: (1) an explicit account of what makes two phenomena comparable (the *tertium comparationis*), and (2) an engagement with processes of falsification designed to rigorously test this assumed comparability (equivalence). These concepts of the *tertium comparationis*, equivalence, and sameness allowed contrastive studies to shake off the critiques of the past by ensuring that researchers compare like with like and that any claims of correspondence are empirically justified, rather than assumed, or worse, established through poor research design.

Key Concepts and Data in Contrastive Studies

The *tertium comparationis* (TC) refers to the assumed comparable common ground that makes comparison possible in a contrastive study. The TC specifies exactly that which is being compared across languages (Connor & Moreno, 2005). For example, the TC may be form-based (e.g., a grammatical category), semantic (e.g., a shared meaning), functional (e.g., a communicative purpose), or discursive (e.g., a genre feature). Crucially, as the TC determines what can be reasonably compared, this determination of comparability must also occur at the level of data design, where comparability must be established across all facets of the texts being compared (e.g., comparability in terms of genre, audience, time period, *inter alia*). However, while comparability is the aim, the ways in which this notion is conceptualized can vary. For example, in comparing two collections of texts, one may argue that the texts should be contemporaneous to ensure comparability. Alternatively, one may hold the view that the texts should be produced in comparable contexts (e.g., under similar economic conditions) and to ensure comparability in this regard, it may be necessary to compare texts that have been produced at very different times. Thus what we deem comparable is guided by our research questions and objectives. Once established, it is notable that the TC always remains a hypothesis. As such, we should see a TC as a working assumption that comparable phenomena exist across languages and the texts being studied. In a contrastive study, this assumption must be tested (Chesterman, 1998).

This process is achieved through equivalence testing. If the TC defines what is assumed to be comparable, equivalence provides the criteria by which that assumption is evaluated. Equivalences operationalize similarity through measurable or interpretable criteria, such as frequency, distribution, semantic meaning, and pragmatic use, *inter alia*. This notion has been referred to in a number of different ways (e.g., correspondence, congruence, equivalent, and equivalence) but the central idea remains consistent. Equivalence testing specifies how an assumed similarity is to be tested and (dis)proven. Krzeszowski (1990) distinguished several types of equivalence, including statistical, translation, system, semanto-syntactic, and pragmatic forms. This was an important contribution, building on James (1980) who proposed a narrower, form-function constraint for equivalence testing. Chesterman's

(1998) contrastive functional analysis approach extended this work further, offering a functionally-driven perspective on equivalence and a rigorous method underpinned by processes of falsification.

These developments allowed contrastive studies to overcome critiques of methodological circularity, as an assumption of similarity (e.g., a shared form) is now tested in terms of criteria such as frequency (does it occur the same number of times in the languages, texts, and contexts being studied?), distribution (does it occur in the same range of texts in the languages and contexts being studied?), meaning (does it share the same meaning in the languages, texts, and contexts being studied?), and so on. There is no limit to equivalence testing, and the more exhaustive the testing, the more critical and accurate any claim of correspondence will be. The only true constraint in equivalence testing is feasibility, as, theoretically, through equivalence testing, the assumed TC transforms into a new, more nuanced TC that then becomes assumed and liable for further testing. However, at some point and for practical reasons, we must cease our investigations and report on our established working theory.

The result of an interrogation of a TC with equivalence testing is a claim about how similar or different the item being analyzed is. Any form of absolute sameness across languages is generally regarded as unattainable, given differences in history, culture, and structure, among others. Even words like 'warm' are culturally and geographically situated, as being warm in the north of England and being warm in Seville are very different things. Thus, in contrastive studies, similarity is treated as a gradient. [Adamska-Sałaciak \(2013\)](#) and [Oleksy \(1986\)](#), among others, distinguish between manifest similarities (i.e., any observable correspondences revealed by an analysis) and maximally similar correspondences (i.e., those correspondences that overlap across the greatest number of equivalences). This graded perspective avoids binary thinking and recognizes that languages may converge in some respects while diverging in others. The question of sameness is further complicated by research design, as convergent analyses, which begin with comparable features identified in two or more languages that are to be systematically compared, will determine sameness based on predetermined shared features. Conversely, divergent analyses, which begin with a feature identified in one language for which a correspondence is searched in others, will allow sameness (or lack thereof) to emerge. Thus, given the evident complexity involved in comparing languages, any contrastive study requires careful design and a critical mitigation of the findings based on that design.

These theoretical developments have been complemented by methodological advances in corpus linguistics and in the development of multilingual corpora, specifically. There are two main types of multilingual corpus. Parallel corpora align source texts with their translations, allowing researchers to trace direct correspondences across languages. This type of corpus is central to research at the intersection of corpus linguistics and translation studies ([Curry & McEneary, 2024](#)) and offers a strong basis for investigating how concepts are translated across languages, contexts, and cultures. Notably, the TC in such corpora is readily established by the fact that the source texts and their translations are intrinsically linked. Comparable corpora, by contrast, consist of non-translated texts collected according to a shared sampling frame. Arguably, these better capture natural language use but they are more fraught as they require careful design to ensure that texts are genuinely comparable. In this way, the TC is not assured, and must be established at each stratum of the corpus. Each type of multilingual corpus has advantages and limitations, and many studies combine both. Ultimately, the choice of corpus will be determined by the research questions that we wish to answer.

Over the last 80 years, contrastive linguistics has undergone a series of transformations and its scope has broadened considerably. For example, building on its traditional foundations, contrastive linguistics can shed light on issues of lexis and grammar ([Aijmer & Hasselgård, 2015](#)), revealing how closely related languages may nevertheless diverge in subtle ways. Equally, contrastive work has shown how the communicative facets of texts (e.g., metadiscourse) can vary across languages, highlighting potential cultural differences in text construction ([Curry, 2024](#)). The field remains closely connected to language pedagogy, where contrastive insights are noted for their capacity to support the raising of plurilingual awareness, intercultural competence, and specialized language teaching in multilingual settings ([Curry, 2022](#)). Its applications have expanded into wider areas of linguistics too, including corpus-based discourse analysis ([Vessey, 2013](#); [Taylor & del Fante, 2020](#)).

[Vessey \(2013\)](#) discusses the affordances of contrastive corpus discourse analysis for offering nuanced insight into discourse practices in multilingual contexts. She contends that to effectively undertake such contrastive work, detailed knowledge of corpus composition, demographics represented therein, and variation within corpus data is needed to avoid false comparisons. In this light, she suggests that researchers should not impose predefined (and what may be ethnocentric) categories on language data and instead allow patterns to emerge through a critical lens. [Taylor and del Fante \(2020\)](#) extend this discussion and propose some solutions to issues of comparison and comparability. In particular, they note that reflexivity is especially important in contrastive work, owing to the complexity of variables shaping the discourses being studied. In essence, these papers demonstrate that in undertaking contrastive discourse work, there is a need to reimagine corpus techniques through a contrastive linguistics lens in order to ensure that comparisons are principled and productive. [Taylor and del Fante \(2020\)](#) note that contrastive corpus and discourse work remains underdeveloped and in recognition of this, in what follows, I present a critical reflection on the affordances of corpus-based contrastive studies for supporting multilingual discourse analysis.

Case Study: Corpus-Based Contrastive Discourse Analyses of Public-Oriented Research Communication

This case study brings together a series of corpus-based contrastive discourse studies that examine how public-oriented research communication constructs knowledge around major societal crises across languages and contexts. Collectively, these studies demonstrate the analytical value of contrastive approaches for unpacking discourse and for comparing how globally shared issues, including the COVID-19 pandemic ([Curry & Pérez-Paredes, 2021](#)), the climate crisis ([Curry, 2024](#)), and the climate and health poly-crisis ([Curry & Brookes, 2025](#)), are discursively constructed in different linguistic and cultural settings. While each study addresses

a distinct thematic focus, they are united by a shared data source (i.e., *The Conversation*, a platform dedicated to disseminating research-based knowledge to the wider public through texts authored by academics and edited by professional journalists), and by a common methodological commitment to principled comparison. In terms of the former, the TC is assured in the composition of each corpus for which all texts come from one source (*The Conversation*), go through the same editorial processes, and have shared genre conventions and communicative goals. This means that any linguistic and cultural variation that is found can be understood as a meaningful insight into how knowledge is discursively constructed in different contexts. In terms of the latter, convergent approaches were employed and equivalence testing involved the interrogation of findings in terms of textual metadata, and wider co-textual and contextual information. Collectively, these studies illustrate how contrastive discourse analyses, underpinned by corpus linguistics approaches, can offer a nuanced perspective on discursive practices involved in knowledge construction across cultures and languages.

One example of this work focuses on the discursive construction of COVID-19 knowledge in English and Spanish. In [Curry and Pérez-Paredes \(2021\)](#), our corpus-based contrastive study examines stance nouns in COVID-19-related academic news blog posts, published in the UK and Spanish editions of *The Conversation*. Using a comparable corpus, the study investigates how authors in each language express stance and attribute sources when communicating on rapidly evolving scientific knowledge. The findings point to broad convergence at the level of genre as, in both languages, the academic news blog posts make extensive use of stance resources to position claims, evaluate evidence, and negotiate uncertainty. However, important divergences emerge. Spanish-language texts show a stronger tendency toward possibility and factuality nouns combined with self-sourcing strategies, which support a more expansive, dialogic positioning. English-language texts, by contrast, favor argument and idea nouns and rely more heavily on external sources, contributing to more contractive rhetorical strategies. These differences suggest culturally contingent preferences in how authority, evidence, and uncertainty are managed when communicating public health knowledge in a shared science-in-progress context.

My second example extends this contrastive approach to climate crisis communication across English, French, and Spanish. In [Curry \(2024\)](#), I analyze the use of direct questions in climate-themed academic news blog posts published by *The Conversation*, examining their frequency, thematic distribution, and metadiscursive and discursive functions across languages. The study shows strong convergence at the level of rhetorical practice, as direct questions are found to be a core feature of public-oriented climate research communication in all three languages. In general, direct questions are found to function as resources for managing text structure, making arguments, and framing hypothetical or future-oriented scenarios. Nevertheless, systematic cross-linguistic differences are also identified. For example, questions are significantly more frequent in the French data when compared to the English data. In addition, the thematic domains in which questions cluster vary across languages, with a focus on science and technology in English, science in French, and politics and society in Spanish. These patterns indicate that while questioning is a shared discursive strategy in climate communication, the issues deemed most worthy of interrogation differ across knowledge cultures, reflecting divergent epistemic priorities.

My third example takes a more inductive approach, examining the discursive framing of the climate and health polycrisis in English, French, and Spanish public-oriented research communication. In [Curry and Brookes \(2025\)](#), we analyze how discourse at the intersection of climate and health is constructed in climate-themed academic news blog posts from *The Conversation*. Specifically, we focus on a dedicated subcorpus of health and climate themed texts and in so doing, we identify the (cultural) keywords: health, *santé*, and *salud*, which form the basis of the study. Notably, this study shows how the careful selection of texts that have a clear shared source for the development of both target and reference corpora can overcome the challenges associated with comparing keywords across languages ([Vessey, 2013](#); [Taylor & del Fante, 2020](#)).

Across all three languages, the analysis reveals convergence in the recognition of climate and health as interconnected issues and in the use of academic news blog posts as a space for contextualizing, explaining, and proposing responses to complex global challenges. However, notable divergences emerge in how this polycrisis is framed. English-language texts tend to emphasize causal explanations and future implications, aligning with empiricist and solution-oriented epistemologies. French texts more frequently engage in definitional and contextual framing, situating the polycrisis as one that connects human, non-human, and plant life. Spanish texts show a tendency to embed climate and health issues within wider transdisciplinary and sociohistorical contexts, foregrounding interconnected crises and structural inequalities. These and the many other differences reported in [Curry and Brookes \(2025\)](#) underscore how shared global problems are refracted through distinct cultural epistemologies.

Considered together, these studies demonstrate the affordances of corpus-based contrastive discourse analysis for unpacking public-oriented research communication. By holding constant the data source (i.e., academic news blog posts published by *The Conversation*), the analyses are able to isolate linguistic and discursive variation that is attributable to language and culture, and wider epistemic contexts. This methodological and theoretical merger of corpus and contrastive linguistics allows areas of correspondence to be identified with confidence, such as academics' shared reliance on stance, questioning, and framing devices to negotiate uncertainty, signal priorities, and shape public knowledge. At the same time, the contrastive analysis brings into relief areas of divergence, revealing how different knowledge-making cultures engage with global unknowns, prioritize different issues, and localize social challenges when communicating about ongoing and unfolding crises.

Conclusion

In this article, I have traced the development of contrastive linguistics from its pedagogical origins to its contemporary realization as a theoretically grounded, corpus-based approach to analyzing language in use. Through this exploration, I have outlined key theoretical and methodological issues in contrastive work and operationalized these notions in a reflective case study, based on a series of contrastive discourse studies of public-oriented research communication. Through this case study, I show that contrastive approaches offer a valuable means of unpacking how shared social challenges are discursively constructed across languages, and, in so doing, I demonstrate how multilingual corpus-assisted discourse studies can reveal areas of convergence and culturally situated divergence in knowledge-making practices. Specifically, the case study illustrates that even within the highly regulated communicative environment that is *The Conversation*, language, culture, and epistemological traditions shape what is foregrounded, questioned, legitimized, or problematized. As global challenges, such as pandemics, climate change, and health crises, increasingly demand coordinated international responses, understanding how knowledge about these issues is framed and constructed across languages becomes ever more critical. Arguably, only with such knowledge in hand can we begin to bridge epistemic gaps and deliver joined-up responses to global crises. In light of this need, expanding multilingual, corpus-based contrastive discourse research will be essential for developing a more inclusive, culturally attuned understanding of how shared challenges are constructed and communicated across diverse linguistic, cultural, and epistemic contexts.

References

- Adamska-Safaciak, A. (2013). Equivalence, synonymy, and sameness of meaning in a bilingual dictionary. *International Journal of Lexicography*, 26(3), 329–345. <https://doi.org/10.1093/ijl/ect016>
- Aijmer, K., & Hasselgård, H. (2015). Cross-linguistic studies at the interface between lexis and grammar. *Nordic Journal of English Studies*, 14(1), 1–8. <https://doi.org/10.35360/njes.337>
- Altenberg, B., & Granger, S. (2002). Recent trends in cross-linguistic lexical studies. In B. Altenberg, & S. Granger (Eds.), *Lexis in contrast: Corpus-based approaches* (pp. 3–48). John Benjamins. <https://doi.org/10.1075/scl.7.04alt>.
- Chesterman, A. (1998). *Contrastive functional analysis*. John Benjamins. <https://doi.org/10.1075/pbns.47>.
- Connor, U. M., & Moreno, A. I. (2005). Tertium comparationis: A vital component in contrastive rhetoric research. In P. Bruthiaux, D. Atkinson, W. Eggington, W. Grabe, & V. Ramanathan (Eds.), *Directions in applied linguistics. Essays in honor of Robert B. Kaplan* (pp. 153–164). Multilingual Matters. <https://doi.org/10.21832/9781853598500-015>.
- Corder, S. P. (1981). *Error analysis and interlanguage*. Oxford University Press.
- Curry, N. (2022). On contrastive analysis and language pedagogy: Reimagining applications for contemporary English language teaching. In L. McCallum (Ed.), *English language teaching in the European Union: Policy and practice across the European Union* (pp. 239–256). Springer. https://doi.org/10.1007/978-981-19-2152-0_14.
- Curry, N. (2024). Questioning the climate crisis: A contrastive analysis of parascientific discourses. *Nordic Journal of English Studies*, 23(2), 235–267. <https://doi.org/10.35360/njes.v23i2.39190>
- Curry, N., & Brookes, G. (2025). The discursive framing of the climate and health polycrisis in English, French and Spanish. In T. Parnell, T. Van Hout, & D. del Fante (Eds.), *Critical approaches to polycrisis: Discourses of conflict, migration, risk and climate* (pp. 309–334). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-76966-5_12.
- Curry, N., & Pérez-Paredes, P. (2021). Stance nouns in COVID-19 related blog posts: A contrastive analysis of blog posts published in *The Conversation* in Spain and the UK. *International Journal of Corpus Linguistics*, 26(4), 469–497. <https://doi.org/10.1075/ijcl.21080.cur>
- Curry, N., & McEnery, T. (2024). The past, present, and future of corpus and translation studies. In D. Li, & J. Corbett (Eds.), *The Routledge handbook of corpus translation studies* (pp. 1–8). Routledge. <https://doi.org/10.4324/9781003057604-7>.
- Fries, C. C. (1945). *Teaching and learning English as a foreign language*. University of Michigan Press.
- James, C. (1980). *Contrastive analysis*. Longman.
- Johansson, S. (2012). Cross-linguistic perspectives. In M. Kytö (Ed.), *English corpus linguistics: Crossing paths* (pp. 45–68). Rodopi.
- Klein, W. (1986). *Second language acquisition*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511815058>.
- Krzyszowski, T. P. (1990). *Contrasting languages*. De Gruyter Mouton. <https://doi.org/10.1515/9783110860146>.
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. Ann Arbor: University of Michigan Press.
- Oleksy, W. (1986). Some recent approaches to equivalence in contrastive studies. In J. Fisiak, D. Kastovsky, & A. J. Szwedek (Eds.), *Linguistics across historical and geographical boundaries: In honour of Jacek Fisiak on the occasion of his fiftieth birthday* (pp. 1405–1420). De Gruyter.
- Taylor, C., & del Fante, D. (2020). Comparing across languages in corpus and discourse analysis: Some issues and approaches. *Meta*, 65(1), 29–50. <https://doi.org/10.7202/1073635ar>
- Vessey, R. (2013). Challenges in cross-linguistic corpus-assisted discourse studies. *Corpora*, 8(1), 1–26. <https://doi.org/10.3366/cor.2013.0032>