Zeitschrift für
SPRACHVARIATION
& SOZIOLINGUISTIK
JOURNAL OF LANGUAGE
VARIATION AND SOCIOLINGUISTICS

https://doi.org/10.5282/jlvs/15

Discourse figures in the Luxembourg language debate (2015–2020)

Christoph Purschke[1]

[1] Department of Humanities, University of Luxembourg, Esch-sur-Alzette, Luxembourg
+
E-mail: christoph.purschke@uni.lu

Received 27.01.2025
Accepted for publication 23.09.2025
Published 24.10.2025

Abstract

The societal role of languages in issues of social integration and collective identity has long been a topic of public debate in Luxembourg. In this paper, I examine the latest wave of this language-ideological debate by focusing on the years 2015–2020. The study uses data from different sources and combines different methodological approaches to examine the contributions and perspectives of various social groups of actors to the structure and dynamics of this debate. I introduce the concept of *discourse figure*, that is, frequent couplings of linguistic forms and ideological motifs tied to different types of actors in discourse, as a unifying analytical lens to the multi-method and multi-sited study of public discourse. The case study is carried out in three steps, first, via a manual analysis of a large media corpus, then, through the computational modeling of a large corpus of user comments from RTL.lu, and last, using a perception study asking participants to make drawings of their experience with multilingualism in Luxembourg. The results show that there is a limited number of discourse figures with structural relevance for the Luxembourg language debate. These figures are found in all datasets and can be classified into three types: language-related, actor-related and feature-related. The ideological core of the central figures in Luxembourg's public discourse reveals a close connection of the discussion on multilingualism and the role of Luxembourgish with questions of collective identity and social integration. From a methodological perspective, the study demonstrates that a multi-method approach – with the help of the concept of discourse figure – can be used to comprehensively reconstruct the dynamics of the Luxembourg language-ideological debate.

Keywords: discourse analysis, Luxembourg, language ideologies, mixed methods, multilingualism

1 Introduction

Public discourses on language are often characterized by multiple parallel dynamics depending on the socio-linguistic makeup of a society, as well as on the groups of actors involved, their ideologic orientations and social positionings. Discourse analysis has developed different methodological solutions to study such complex debates at the intersection of corpus linguistics and sociolinguistics (Spitzmüller & Warnke 2011;

Tannen et al. 2015): Some approaches combine statistical methods of corpus analysis with manual pattern interpretation, for example, in Corpus-Assisted Discourse Studies (CADS; Ancarmo 2020; Vessey 2017); studies in Multimodal Discourse Analysis (MDA; Kress & van Leeuwen 2001) focus on the interaction of multiple modes of communication; Androutsopoulos (2024) uses a scaled mixed-methods approach to the study of visual cues in discourse; and work in Anthropological Discourse Analysis (ADA; Philips 2020) collects data from different types and multiple sites of interaction.

This paper is inspired by such approaches but derives its main interest from two circumstances: the particular structure of the Luxembourg language debate and an orientation towards multi-method and interdisciplinary approaches in discourse studies. The guiding assumption is that the discourse in Luxembourg requires the analysis of different sites and modalities of discourse depending on the group of actors involved, and that the social positionings of each group call for different methods of data collection and analysis. The overarching research question is therefore twofold:

1. Which discourse figures, that is, couplings of linguistic patterns and ideological motifs used by actors in discourse, can be identified for different groups of actors?
2. What contribution does the methodological setup make to a comprehensive analysis of language ideological debates (Blommaert 1999)?

Using the example of the public debate on multilingualism and identity in Luxembourg, in particular the years 2015–2020, I carry out a multi-method discourse analysis focusing on central figures in discourse to examine the discussion and its societal anchoring from different angles. In doing so, I partly draw on existing studies but reinterpret their data against the backdrop of the research question. Such a multi-angled approach seems especially relevant as this debate was characterized by structural peculiarities revolving around the special role of Luxembourgish as a national language, and embedded in a language regime that is dynamic in several ways. First, the country's complex multilingual makeup is shifting between the poles of demographic dynamics and national identity. Second, there is a state-led language policy development of multilingualism, particularly Luxembourgish. Third, the role of individual languages became the subject of a heated debate between 2015 and 2020, particularly because of a national referendum on the right to vote for foreign residents in 2015 and the following populist and political instrumentalization of the issue (Garcia 2014; Rivera Cosme 2023).

Against this backdrop, Section 2 outlines the sociolinguistic background of the case study concerning its socio-demographic, political and ideological foundations. Section 3 then introduces my multi-method approach by compiling a set of methods aiming to reconstruct the perspectives of different groups of actors in this discourse. Subsequently, data analysis in Section 5 puts the idea of *discourse figures* (Section 4) to an empirical test. I trace typical linguistic patterns and ideological motifs across three datasets while focusing on different social groups present and active in this discourse. In this way, the study aims at reconstructing the socio-pragmatic complexity of the Luxembourg language debate. Finally, Section 6 discusses the potential and pitfalls of my approach, including a critical evaluation of the term *discourse figure*, as well as overarching findings from the individual studies.

2 The case study: Luxembourg

Luxembourg is characterized by a historically grown, complex societal multilingualism (Erhart & Fehlen 2011). The country has three official languages enshrined in law: Luxembourgish as the national language, and German and French as administrative languages. Multilingual practice is organized along social domains, with different languages taking the dominant role depending on the domain. German, for example, is the traditional language of school literacy and the (print) media, while French serves as the language of legislation and is dominant in private business contexts. Luxembourgish is primarily used among Luxembourgers but also plays an important role in institutional communications and for social integration of foreign residents. Additionally, Portuguese and Italian are crucial as minority languages. Still, English is becoming increasingly important as a lingua franca for the rapidly growing proportion of foreign residents (47.7% of 660k inhabitants; STATEC 2023).

Research into Luxembourg and its multilingual makeup has recently seen a rise in interest (see Purschke & Gilles 2023 for an overview). The country's language regime is currently dynamic in several respects. Luxembourgish is undergoing societal and political development to become a fully-fledged standard language (Gilles 2023). This development runs parallel to the pragmatic domain expansion of written Luxembourgish (Gilles 2015), which was largely driven by the advent of social and digital media. The increasing use of Luxembourgish in public communications, the sharp rise in demand for Luxembourgish as a foreign language courses and the politically driven development of orthography, lexicon and university teaching are evidence of a readjustment of the position of Luxembourgish in the language regime.

This is accompanied by a societal revaluation of the language (Garcia 2018), which was traditionally seen primarily as a spoken variety among Luxembourgers, and linguistically as a dialect of German. At the same time, the institutional anchoring of Luxembourgish is only partially complete. It is barely present in school curricula, so the population does not have comprehensive (meta) knowledge of spelling or grammar. In addition, Luxembourgish is under pressure due to its close structural and socio-pragmatic contact with German and French, which is reflected, for example, in a lexical and attitudinal orientation toward German among younger speakers (Conrad 2017).

The changing role of Luxembourgish in the fabric of multilingualism is a critical factor in the current dynamics. It is also evidence of the ideological tensions that characterize the country's public discourse on multilingualism and that are reflected, for example, in the uncertainty of young speakers when it comes to situating Luxembourgish (Bellamy & Horner 2018). The economic and socio-demographic development of the country also contributes significantly to this situation, for example, through the high number of daily cross-border commuters (227.7k; STATEC 2023) and the associated practical requirements of a multilingual society in which many employees have no or only limited access to Luxembourgish (or, in many cases, any of the other official languages). This leads to a situation in which Luxembourgish is recording increasing numbers of speakers and at the same time becoming increasingly marginalized (Fehlen et al. 2023).

Against this backdrop, the years between 2015 and 2020 saw a new "wave" in the ongoing public debate on the status of Luxembourgish in the country's multilingualism and on the role of the official languages for constructing a collective identity (see Péporté et al. 2010 and Fehlen 2016 on the history of the debate). The national referendum held in 2015 on voting rights (at the national level) for foreign residents, which ended with a clear rejection of the proposal, has been identified as the starting point of this debate (Purschke 2023). In this context, a small group of activists (Nee 2015/Wee 2050) formed around geography teacher – and current member of parliament for the right-wing populist ADR – Fred Keup. This group spoke out strongly against the right to vote for foreigners, instrumentalizing the societal role of Luxembourgish for an anti-growth (and latently xenophobic) campaign. On their website and Facebook group page, they also proclaimed to act as the mouthpiece of the societal majority, that is, for the 80% who voted agaist the referendum. In 2016, two public petitions were then discussed in parliament, which dealt with opposing language policy issues: petition 698, which aimed to upgrade Luxembourgish to the primary official language, and petition 725, arguing in favor of maintaining the multilingual status quo.[1]

These topics were widely – and often uncritically – picked up by the national news outlets and in social media debates. As a consequence, and supported by the very active Facebook page of the Nee 2015/Wee 2050 group, a discourse climate emerged in which the role of Luxembourgish was stylized as a question of national self-assertion. This development was driven mainly by the 2018 election campaign, that is, the political parties relied on folkloristic and identity-related motifs in response to the supposed pressure from the population, with Luxembourgish as the central ideological vehicle. In addition, at the instigation of the then Minister of Education, Claude Meisch, a series of language policy initiatives was launched to strengthen the status of Luxembourgish: In 2017, a 20-year *Strategie fir d'Promotioun vun der Lëtzebuerger Sprooch* ('Strategy for the promotion of the Luxembourgish language') was presented, from which a new

language promotion law emerged in 2018. This law provides for a whole range of structural support measures, including a commissioner and a ministry-led center for the Luxembourgish language, as well as new initiatives for its embedding in school and university curricula. In addition, an updated version of the orthography was introduced in 2019.

One striking aspect of the public debate was that the issues surrounding Luxembourgish and multilingualism discussed in the media and in politics seemed to play the role of a proxy discourse, in which pressing societal issues such as demographic and economic development were invoked but only partially visible (Purschke 2023). Instead of discussing pressing societal issues like social justice, mobility concepts or affordable housing, people talked about language and identity. In this context, the language issue was primarily a vehicle rather than the true subject, with the help of which topics such as fear of cultural marginalization and economic decline could be anchored in discourse. Additionally, the public discussion in part appeared to be a sham debate that a small group of actors artificially inflated to produce news and entrench populist stances in the discourse. Against this backdrop, I take a closer look at the structure, recurring linguistic patterns and ideological motifs of the language debate from a multi-method perspective, with a particuar focus on different groups of social actors.

3 A multi-method approach to discourse analysis

This study advocates for interdisciplinary work and the use of mixed-methods approaches. In sociolinguistics, this has been implemented in different ways, for example, by comparing corpus-based and qualitative findings on the same topic (Baker & Levon 2015), by contrasting collocations with word embeddings for corpus analysis (Batchelor 2024) or combining different methods for the same purpose (Kircher & Hawkey 2022; Purschke 2020a). In recent years, there has been a growing interest in connecting sociolinguistic and computational methods under the label of Computational Sociolinguistics (CSLX; Grieve et al. 2023). Discussions on the use of mixed methods often revolve around the combination of quantitative and qualitative approaches, drawing on theoretical (macro-scale vs. in-depth modeling), methodological (quantification vs. interpretation) and analytical (explanatory vs. hypothesis-driven analysis) differences. For this study, I am less interested in the differences between those types of methods than in the ways they are used to establish categorizations for analysis, that is, variable-based holistically (quantitative default) and case-based selectively (qualitative default; Borgstede & Scholz 2021). Rather than distinguishing between qualitative and quantitative methods for conceptual reasons, the choice of methods for my study follows the specificities of the discourse under investigation as well as a purpose-oriented understanding of methods.

I therefore choose methods based on their analytical contribution to the overarching research question – and thus indirectly based on the specific structure of the Luxembourg language debate. In doing so, I identify three axes of structural differentiation and three groups of social actors in discourse, which set the basis for the methodological setup:

- *Official vs. private functions:* I distinguish between actors with official functions in discourse, that is institutional, political, journalistic or scientific duties, as opposed to private actors.
- *Insider vs. outsider positions:* In many respects, the Luxembourg public sphere is divided into two groups, with a distinction being made between insiders (autochthonous Luxembourgers) and outsiders (foreign residents, that is, (non-)naturalized residents with migratory background, expats, cross-border workers). I am aware that the outsider group comprises different kinds of actors depending on their relation to the country (residents vs. non-residents, foreign vs. naturalized residents, expats vs. migrant workers). The relevant distinction for the language debate, however, holds despite such sub-differentiations.
- *Active vs. passive roles:* In addition, I distinguish between actors that can participate actively in discourse (that is, mostly Luxembourgers and holders of social functions) and those who are unable to do so, for example, because they lack the necessary language skills or a legitimate

position as a speaker – which in many cases overlaps with the outsider category.

Using these criteria, different aspects of the discourse can now be methodically targeted, whereby each sub-group requires the choice of suitable methods depending on their respective societal anchoring. This is also linked to a difference in the media in which individual actors can express themselves, for example, party manifestos and news outlets as media of official communications vs. social media chats and user comments for private actors. Following this rationale, I choose the following methods a), b) and c) to examine crucial sections of the Luxembourg language debate.

a) Media corpus analysis: Since 2015, I have collected publicly accessible statements by Luxembourg actors to document the discourse's development. The texts come primarily from public media (newspapers, television, web), government websites, book publications, election campaign communications and the social networks central to Luxembourg (Facebook, Instagram, X). The corpus comprises roughly 1000 entities that were manually cataloged according to the above criteria and searched for recurring linguistic patterns, ideological motifs and actors involved (Purschke 2023). In contrast to CADS studies, I do not use corpus quantification for data exploration and, hence, do not report on frequencies for patterns found, but analyze the corpus manually looking for recurring patterns. In doing so, I adopt an ethnographic approach to corpus analysis (Tusting 2020) that acknowledges the respective epistemological value of corpus linguistic and ethnographic approaches to corpora (Hodge & Goico 2022) and reflects my own position in the discourse in question, that is, a naturalized Luxembourger of German descent with a social function as an academic researcher and an active role in the language debate.

One reason for this choice of method relates to the fact that the collected statements cover a variety of different media, modalities and languages, rendering structured corpus building rather complicated without extensive preprocessing (like text extraction from videos, image descriptions or text translations). Additionally, with Luxembourgish still in the process of standardization, its written form exhibits an abundance of spelling variation, even in formal communications. This makes it difficult to perform procedures such as co-occurrence analysis on this data, even with prior text normalization (Lutgen et al. 2025). For this study, I restrict the corpus to statements by official actors who also occupy the position of insiders and have active roles in the discourse.

b) Computational corpus modeling: In a second step, I look at the statements of insiders who play an active role in the discourse but (typically) do not hold any official functions. The largest available data source for this is user comments on journalistic content on RTL.lu, the central news portal in Luxembourg. The service has played a vital role in developing written Luxembourgish for many Luxembourgers, both in terms of reading (journalistic content) and writing (user comments). I have access to RTL's entire text archive, which at the time of the study included 585,358 (anonymous) user comments and 179,298 news articles from 2008 to 2018. The amount of data requires a computational approach for modeling and analysis. I preprocessed the data to reduce the amount of orthographic variation in the texts (Purschke 2020b). Next, I trained a word embedding model (Word2Vec; Mikolov et al. 2013) on the user comments. In the model's high-dimensional vector space, language data is represented according to contextual similarity so that words that appear more frequently in comparable pragmatic contexts, are realized in similar syntactic structures or have a similar meaning are assigned higher similarity values, that is, the cosine distance between their vector representations in the model. Research has shown that relationships between the vectors representing linguistic concepts in embedding spaces can be meaningfully interpreted, for example, regarding semantic relations between concepts (Garg et al. 2018) and semantic change over time (Kozlowski et al. 2019), or used for discourse analysis (Bubenhofer 2022).

In this study, I read the word vectors' contextual similarity in the embedding model as claims for the relative discursive closeness of the concepts in the data. In doing so, I understand this relationship holistically, resulting from a confluence of socio-pragmatic elements ingrained in language use that reflect a word's relative position to others in the corpus and, consequently, in this section of public discourse. Based on the results of the media corpus analysis, I submit search

queries to the model for central discourse figures and analyze the ten closest word vectors to these to reconstruct their discursive context. Since the texts are almost exclusively in Luxembourgish, this study focuses on the positionings of insiders without official functions who can, however, play an active role in the discourse via the medium of user commentary.

c) Perceptual mapping study: To counter the insider perspectives with an outside view, I then look at the positionings of actors with an outsider position to public discourse in the country, in this case, foreign residents. Since these usually do not play an active role in the discourse and only rarely assume official functions, I need a different method for this. I choose a perceptual mapping experiment to capture individual experiences with the multilingual public sphere in Luxembourg (Purschke & Schmalz 2022). This method uses blank drawing canvases to capture the participants' conceptualizations of the object of investigation in any graphical style they prefer. I used only verbal instructions ("How do you experience multilingualism in Luxembourg? Make a drawing.") for the drawing task and collected drawings in a two-step process: 1. individual drawings followed by an interview on drawing strategies and central motifs; 2. collective drawings by two participants with a follow-up interview on the negotiation of visual and content-related aspects. I collected 28 data points consisting of 56 individual and 28 collective drawings that stem from a seminar project conducted in 2017. The data are socio-demographically broad containing almost exclusively contributions by foreign residents. I analyze the data for dominant types (everyday situations, mind maps, geographic projections), perspectives (personal, group-based, country-based), design strategies (simple vs. complex, detailed vs. reduced) and thematic motifs. For this study, I focus on the motifs relevant to the research questions.

4 Discourse figures

In this paper, I am particularly interested in the connection between social positioning strategies of actors in the language debate, the ideological motifs used to mark one's own position in interaction and the linguistic means carrying these positionings. In doing so, I take a theoretical starting point from discourse analysis – with a special interest in collective public practices of knowledge production (Keller 2011), theories of stance (Jaffe 2009) and metapragmatic positionings towards topics and actors (Spitzmüller 2023), aiming at a holistic understanding of how people organize everyday practice through negotiating social belonging via language (Purschke 2020a).

Highlighting the close connection between a linguistic form, its ideological grounding and the person responsible for a statement, I introduce the concept of *discourse figure* to capture the fact that prominent stances in discourse are not only recurring couplings of linguistic patterns and ideological motifs but are tied to (types of) actors. That is, the same stance might invoke different social implications depending on the social position (e.g., insider vs. outsider), function (e.g., official representative vs. private citizen) and role (active vs. passive participation) in the debate. In this sense, the concept highlights the connection between stance-taking and a socially grounded theory of action (Schwarz 2021). I understand discourse figures as socio-pragmatic figure-ground relations (Wertheimer 1925), that is, as the relationship between a socio-cultural ground (discourse) and a socio-pragmatic stance (figure). Discourse figures highlight ideological motifs by using topical vehicles in a specific linguistic form to broker social relations against the background of a discourse, carried out by particular actors. To this effect, the term describes couplings of linguistic forms and ideological motifs in public statements by actors when used to address a particular topic and position socially.

The main reason to do so relates to the fact that acts of social positioning through language do not exist outside of but are determined, reproduced and upheld by complex social structures in society. The question of who can circulate which kind of stance, with what impact and social consequences (both for the actors and the target audience) is of vital importance for discourse figures. For example, advocating for the recognition of Luxembourgish as an official language of the EU in the sentence *Lëtzebuergesch muss endlech eng offiziell EU-Sprooch ginn!* ('Luxembourgish must finally become an official EU language!') can be described in terms of its ideological grounding and linguistic form, but depending on the actor taking this

stance (e.g., the Minister of Culture vs. an anonymous user on RTL.lu), it might be anchored differently in the discourse (e.g., as authoritative vs. a form of protest from below), circulated differently (e.g., as an official policy vs. a private agenda) and lead to different consequences (e.g., an official request for recognition with the EU vs. a salty reply by another user on RTL). In the following, I use the concept of discourse figure as a unified analytical lens for the study of the Luxembourg language debate.

5 Tracing discourse figures

In this section, I report on the three different sites of the debate that capture central groups of actors and their social positionings in this discourse as explained above. In doing so, I trace central discourse figures across the different datasets and methods to compare their presence and pragmatic pertinence for the case study. While I limit my study to the Luxembourg case, I am aware that similar cases have been documented in the literature (Blommaert 2011; Naglo 2007; Watts 1999), however, without adopting a multi-method approach.

5.1 Media corpus analysis

For the first study, I analyze the collected media corpus on the language debate in Luxembourg focusing on actors in insider positions with active roles in discourse, who often additionally assume official functions in the public sphere. These include above all politicians, researchers and journalists, but also the aforementioned Facebook group of language activists. Impulses from this small group of actors were often taken up by citizens in comment forums (user comments on RTL.lu) and social networks (mostly Facebook), and thus anchored in discourse, often reflected in new uptake by national news outlets and increased circulation across different media. In doing so, I am aware that, in the age of social media and the blurring of producer/consumer logics in digital discourses, such a hierarchical organization of circulation is not necessarily always plausible. However, the specific structure of the Luxembourg public sphere, with its strong orientation towards state and para-state institutions and the absence of a broad critical public, validates this approach.

I find a limited number of discourse figures, which occur repeatedly across different contexts, media types and groups of actors, which were thus assigned the role of ideological leitmotifs (Purschke 2023). This includes, for example, the juxtaposition of Luxembourgish and multilingualism in the context of language-political statements or the assessment of traditional Luxembourgish features as 'authentic' or 'honorable' in opposition to French and German loan structures that are sanctioned as 'foreign' or 'imported'. I summarize the found discourse figures into three emblematic types (language-related, speaker-related, feature-related) based on their linguistic structure, the embedded ideological motif and the "non-obvious meanings" (Partington et al. 2013: 11) they contain and discuss examples of all types in the following. That is, all three types of discourse figure can be characterized by a complex semantic coding, that is, they combine an explicit ideological motif with an underlying – often contrary – second.

An example for the first type of figure is frequently found in statements by politicians, particularly during the 2018 parliamentary election campaign, for example, by Minister of Education Claude Meisch: 'Luxembourgish and multilingualism can be promoted simultaneously'. On the surface, this figure integrates two socio-symbolic poles in the Luxembourg language regime, that is, the practical requirements of multilingualism in a highly diverse society and the need for cultural representation through Luxembourgish as a symbol of a national identity. It also performs an ideological double address to the advocates of multilingualism and the representatives of a Luxembourgish-first ideology (Horner & Weber 2008). At the same time, however, it does the opposite regarding its ideological grounding by constructing multilingualism and Luxembourgish as a rhetorical opposition outside actual practice and, hence, by sustaining the ideological opposition between the national language Luxembourgish and the "other" official languages. A similar example of the same type would be the 2018 campaign slogan of the liberal party DP *Zukunft op Lëtzebuergesch* that can be read either as 'A Luxembourgish version of the future' (referring to the country) or as 'A future in Luxembourgish' (referring to the language). In this manner, this type of discourse

figure traverses the ideological landscape of tension that is characteristic of the Luxembourg public sphere. On the one hand, it frames Luxembourgish as a constituent of national identity, while on the other hand, it thematizes the practical requirements of a multilingual society shaped by labor migration and cultural diversity using languages as the thematic vehicle.

The second type of discourse figure is characterized by a similar structure; however, it is not languages that serve as vehicles, but rather the focus is on the speakers and their legitimate position in the discourse. An example of this type can be found in the ubiquitous designation of Luxembourgish as *eis Sprooch* ('our language'). This type of figure can be found in all communicative contexts and among all types of actors. It serves, above all, to characterize the role of Luxembourgish in the country's multilingualism as unique compared to German and French. The genesis of this figure can be traced to the nation-building process in Luxembourg that was fraught with conflict during the 19th and early 20th centuries. During this period, Luxembourgish assumed a pivotal role as a marker of a distinct cultural identity (Péporté et al. 2010). The figure is also of central discursive importance for the second half of the 20th century, that is, for the political development of Luxembourgish to a national language before 1984 (Fehlen 2016). From a historical perspective, this figure resonates with the socio-cultural differentiation of the Luxembourgers from the neighboring language areas – and indirectly powers –, which were formative for the present shape of the country and its official trilingualism.

The current function of this figure in the discourse is also characterized by a delimiting ideological motif that establishes a linguistic contrast between Luxembourgish and the other two official languages. What is more, the figure is inscribed with an ideological stance of belonging through legitimate linguistic competence (Bourdieu 1982), in that it identifies Luxembourgish as the property of Luxembourgers via the possessive pronoun. At the same time, the figure implicitly serves an integrating function that refers to the historical role of language in Luxembourg's creation and the importance of Luxembourgish as part of a collective identity. However, this integration is linked to specific prerequisites, above all, the acquisition of legitimate language skills as proof of belonging. Consequently, this excludes all those to whom Luxembourgish is not accessible – even though they practically participate in the community –, as can be seen in similar claims often heard by politicians of the right-wing populist ADR, such as Fernand Kartheiser: 'Multilingualism is a fantastic asset. But when people come to Luxembourg to work, they should learn and use Luxembourgish.'

This kind of socially excluding motif becomes particularly clear in the third type of discourse figure, the shibboleth (Purschke 2014; Gumperz 1982), that is, figures that instrumentalize the language use of a particular group that deviates from a situational norm to label these speakers as foreign and endorse their social sanctioning. Often, these figures are characteristic of a whole complex of everyday situations and the language-ideological conflicts enshrined in them. In Luxembourg, this type includes the proverbial *En français, s'il vous plaît!* ('In French, please!') that is representative of many everyday situations in which (Luxembourgish-speaking) Luxembourgers encounter non-Luxembourgish-speaking residents or cross-border commuters. These include, above all, conversations in stores, whereby <shopping at the bakery> has become a leitmotif of the discourse. But the healthcare sector, being largely dependent on foreign-born professionals, is also exemplary of this figure.

Its function is primarily to highlight the supposably excessive presence of French (proxying to "French-speaking people") in the public sphere, often combined with statements about the apparent decline of Luxembourgish. In this way, it also serves the purpose of countering statements from politicians and the media, which typically assert the seamless functioning of multilingualism, with an account of "actual" language practice, thus legitimizing the claim that Luxembourgish is threatened by the (too) strong presence of French – and increasingly disappearing from the public sphere. This figure is frequently linked to the allegation of an absence of willingness to (linguistically) integrate. That is, migrant workers are expected to adapt to the Luxembourg language market, where Luxembourgish is stipulated as the normative *lingua franca* by actors who circulate this figure in discourse as a way to ostracize those groups of speakers. Similar figures can be found

in the corpus from conservative and right-wing politicians, such as: 'When I go into a store or hospital, I don't want to speak a foreign language. I feel at home here and want to speak my own language.'

Regarding the research question, the study indicates that manual corpus analysis can be employed to identify central types of discourse figures and examine them in terms of their linguistic and ideological structure, especially in cases where the data base is too varied (medially, modally, linguistically) to apply corpus statistical approaches. The method allows for a context-sensitive and fine-grained analysis of statements by official actors with insider positions and active roles in shaping the discourse. In a second step, I compare the results of the media corpus analysis with the computational modeling of a large corpus of user comments to examine the perspectives of insiders to the discourse that participate actively, yet not in an official capacity.

5.2 Computational corpus modelling

For this study, I am interested in the individual statements by readers of the RTL.lu news platform, which they leave in the form of user comments under journalistic articles, and which frequently deal with the topics of multilingualism and social belonging. To test this, I perform search queries for the most contextually similar words in the embedding model using examples of the three types of discourse figures as keywords. In this way, I aim to reconstruct socio-pragmatic contexts from their statistical aggregation in the vector space. I assume that the overall distribution of and the distance between word vectors in the embedding model capture meaningful information about a multitude of individual statements, and that lists of words similar to the searched keywords are indicative of recurring linguistic patterns in the dataset.

As a first example, I use the two opposing terms for the Luxembourgish national language: *Lëtzebuergesch* ('Luxembourgish' [official name]) and *eis Sprooch* ('our language' [discourse figure]). To test the pertinence of the figure in the comment data, I also compare the two keywords to the term *Sprooch* ('language') alone. The values behind each word in the list indicate the vector similarity between that entry and the respective keyword, with higher values representing more similar vectors. Additionally, I give grammatical information for entries, with the term [variant] representing spelling variants to lemmas.

- ***Lëtzebuergesch:*** *Franséisch* ('French' N 0.870965), *franséisch* ('French' Adj 0.838781), *Franséich* ('French' N [variant] 0.811675), *Franzéisch* ('French' N [variant] 0.796270), *Englesch* ('English' N 0.770707), *englesch* ('English' Adj 0.725501), *Lëtzbuergesch* ('Luxembourgish' N [variant] 0.701603), *Däitsch* ('German' N 0.690312), *Sprooche* ('languages' N 0.682778), *Letzeburgesch* ('Luxembourgish' N [variant] 0.669334)
- ***Sprooch*****:** *Mammesprooch* ('mother tongue' N 0.814757), *sprooch* ('language' N [variant] 0.771097), *Landessprooch* ('national language' N 0.759516), *Schreifweis* ('spelling' N 0.751803), *Sproch* ('language' N [variant] 0.723642), *Nationalsprooch* ('national language' N 0.723429), *Orthographie* ('orthography' N 0.701390), *Identitéit* ('identity' N 0.692551), *Friemsprooch* ('foreign language' N 0.660245), *Nationalitéit* ('nationality' N 0.656720)
- ***eis + Sprooch:*** *Integratioun* ('integration' N 0.658710), *dSprooch* ('the language' Det + N [variant] 0.649252), *Nationalsprooch* ('national language' N 0.645067), *Sproch* ('language' N [variant] 0.644663), *Identitéit* ('identity' N 0.639937), *ons* ('our' Pron 0.623955), *Traditiounen* ('traditions' N 0.610526), *Mammesprooch* ('mother tongue' N 0.610475), *Franséisch* ('French' N 0.601850), *Däitsch* ('German' N 0.601521)

The comparison of the ten most similar words for *Sprooch* first reveals how discursive closeness works in embedding models, that is, what kind of similarity is represented here. The most similar word vectors include related terms from the same semantic field, such as 'foreign language' and 'national language' or 'spelling', but also a spelling variant with *sprooch* and a popular misspelling of 'language' with *Sproch*.[2] However, 'identity' and 'nationality' already indicate that language is part of a discursive context linked to the discourse surrounding multilingualism and

integration. In contrast, the most similar vectors for *Lëtzebuergesch*, apart from the plural form 'languages', only contain language labels – in several spelling variants. This indicates that the contextual similarity in this case seems to be determined primarily by syntactically similar structures in which people wrote about languages.

Contrary to this, the word vectors for the discourse figure *eis Sprooch* show, next to the expectable variants of the search terms (*sprooch*, *Sproch*, *ons*), that the combination of these words refers to contexts in which people talk not only about language (see the difference to the search query for *Sprooch*) but also about questions of social identity and integration ('traditions', 'national language', 'identity'), and this in the context of the other official languages *Däitsch* and *Franséisch*. The aggregative representation of large amounts of text in embedding models thus seems to capture aspects of discourse pragmatics that can be read as discursive contexts for the identified thematic figures.

To corroborate this further, I query the model with two further figures, the beginning of the shibboleth *En français, s'il vous plaît* and the topical situation <shopping at the bakery> using the terms 'language' and 'baker' / 'bakery' as keywords:

- ***Sprooch + Bäcker:*** *Sproch* ('language' N [variant] 0.691106), *Croissant* (N 0.660578), *Mammesprooch* ('mother tongue' N 0.655153), *Lëtzebuergesch* ('Luxembourgish' N 0.628489), *Franséisch* ('French' N 0.622765), *Vendeuse* ('saleswoman' N 0.620364), *Friemsprooch* ('foreign language' N 0.612616), *Metzler* ('butcher' N 0.610580), *Schräiner* ('carpenter' N 0.598761)
- ***en + français:*** *Francais* ('French' N [variant] 0.682297), *s.v.p.* ('please' [abbreviation] 0.640394), *svp* ('please' [abbreviation] 0.623281), *Français* ('French' N [variant] 0.604420), *Franséisin* ('Frenchwoman' N 0.602703), *plaît* ('pleases' V 3rd singular 0.589179), *Parlez* ('speak' V 2nd plural 0.562329), *s'il* ('if it' 0.553853), *een* ('one' Pron 0.538637), *s.v.pl* ('please' [abbreviation] 0.533539)

Both discourse figures are also represented in the aggregated data. In the case of <shopping at the bakery>, we find a combination of contextually similar terms that refer to a prototypical situation frequently encountered in the media corpus. These are descriptions of the fact that (paraphrased using the terms in the list) 'in a bakery, it is often not possible to buy a croissant in Luxembourgish, that is, in one's mother tongue, because the saleswoman only speaks French, a foreign language.' The fact that the extended discourse figure constituted by these terms is so clearly represented in the embedding model, and this in the user comments, not in the journalistic texts, shows how discourse figures can seep into the population through media uptake and thus assume a central function for the discourse. Another reason might be that this type of situation is in fact part of many people's everyday routines, and, hence, potentially backed by experience.

In the case of the shibboleth figure, the picture is even clearer. The comment corpus was filtered by language before the model training to use only Luxembourgish texts. Nevertheless, given the tendency of many Luxembourgers to engage in situational code-switching, a certain amount of foreign-language material remains in the corpus. This is clearly reflected in the search results. Almost all similar word vectors directly relate to the figure *En français, s'il vous plaît*. Only the indefinite pronoun *een* ('one') indicates that the form *en* is also common in Luxembourgish as a personal pronoun and indefinite article. The shibboleth function of this figure, therefore, seems to be confirmed, although with this list, we do not learn much about its larger discursive context. Comparing this result to a close reading of comments in which the keywords and similar word vectors appear (see Example 1), however, shows that in most cases, these words in fact are used in contexts where users discuss language-related issues such as being required to do shopping in a foreign language.

(1) Dir gitt op eng Visite guidée, déi op Lëtzebuergësch ugekënnegt as. De Guide äntwert 'Quoi?', 'Comment?', 'En français', oder 'En français, s'il vous plaît' (wann et da gutt kënnt!) An da sot emol eppes, da kritt Der vum Guide ze héieren, Dir wäert Rassist, Arborigène,

> Inculte, oder soss eppes. .... an déi “kültivéiert” Lëtzebuerger am Grupp kucken Eech wéi wann der vum Mound kéimt! Dann denkt Der einfach, klibbert mech, a gitt Ärer Wee! Beim Bäcker hutt Der meeschtens keng Chance, wann Der gär eng Aachtchen, eng Schneck oder eng eng Äppeltäsch hätt. Am “Fachjargon” heescht dat ‘un huit’, ‘un escargot’ an une ‘poche aux pommes’. .... oder Dir hongereg heem! An dat an Ärem Land! [RTL corpus, ID 10144]
>
> ‘You go on a guided tour, which is announced in Luxembourgish. The guide answers ‘Quoi?’, ‘Comment?’, ‘En français’, or ‘En français, s’il vous plaît’ (if it goes well!) And then say something, and the guide will call you racist, aboriginal, uncultured, or something else. .... and the “cultured” Luxembourgers in the group look at you like you came from the moon! Then you just think, piss off, and go your way! At the bakery, you usually don’t have a chance if you want an ‘Aachtchen’, a ‘Schneck’ or an ‘Äppeltäsch’. In “professional jargon” this is called ‘un huit’, ‘un escargot’ and une ‘poche aux pommes’. .... or you [go] home hungry! And that in your own country!’

The results of the computational corpus modeling are deemed to be beneficial in regard to the research question. The figures identified in the media corpus analysis can also be found in the aggregated discursive contexts of the vector model. This suggests a close interdependence between the statements of actors with official and private functions in the Luxembourg discourse, which can be read as an indication of the mutual influence of the two groups, also considering the small size of the Luxembourg public sphere. Nevertheless, a comprehensive understanding of the discourse in question and the pertinent keywords to use is essential for a fruitful analysis. Concerning the method, it is worth noting that word embedding models deal better with orthographic variation than traditional corpus linguistic methods and are therefore better suited to be used for Luxembourgish when looking to reconstruct discourse figures in context.

5.3 Perceptual mapping study

So far, my study has only considered the perspectives of actors that have insider positions in the discourse and participate actively in its development. Given the demographic composition of the country, it seems necessary, however, to also examine the other half of the resident population, which, in many cases, is seen to represent an outsider perspective, especially considering expats and migrant workers. Since there are not many outsider voices present in the Luxembourg public sphere, aside from university experts, I need a different methodological approach to gain insight on this section of the discourse.

I opt for a perceptual mapping experiment in combination with participant interviews and examine the 84 drawings for evidence of discourse figures and the discussion about multilingualism in the country (Purschke & Schmalz 2022). I first notice that most drawings paint a positive picture of multilingual practice. Many drawings emphasize the coexistence of different languages in everyday situations or in relation to Luxembourg’s external borders as a container, as well as individual multilingual repertoires. This indicates that focusing on the non-Luxembourgish perspective can add a different angle to the analysis of the language debate, especially considering the predominantly negative tone in the RTL.lu user comments.

A total of 21 drawings in the dataset deal directly with the practical organization of multilingualism, be it in the reproduction of linguistically mixed everyday situations, in drawings that anchor multilingualism

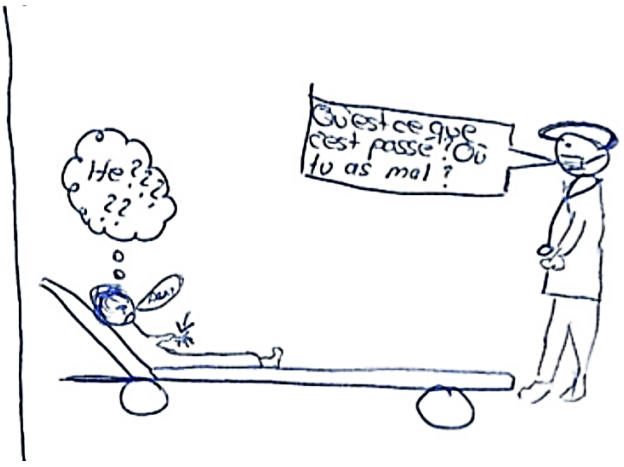

Figure 1: Participant drawing: Prototypical situation <appointment at the doctor>.

Figure 2: Participant drawing: Prototypical situation <shopping at the bakery>.

geographically or in depictions of individual multilingual repertoires situated in everyday life. Of these, three images are particularly revealing because they refer directly to the discourse figures found in the other studies. The first two are reproductions of the two prototypical situations <appointment at the doctor> (Figure 1; Excerpt from the overall drawing) and <shopping at the bakery> (Figure 2). The accompanying interviews show that the participants see these situations as exemplary of contexts in which multilingual practice may reach its limits.

However, there is a difference in the motivation of the two drawings. Visiting the doctor is described in the interview as challenging in practice since, unlike at the bakery, trouble-free communication with the doctor may be a prerequisite for correct treatment. In the drawing (Figure 1), the doctor asks the patient in French about their pain, and the patient signals a lack of understanding (in the thought bubble). Additionally, they utter *Aua*, which, according to the participant interview, signals a lack of ability to describe pain in French accurately. In contrast, the artist of the bakery scene (Figure 2) makes direct reference to the discourse in the interview and characterizes it as a pseudo problem that some Luxembourgers are upset about on purpose. Therefore, in the drawing, the client on the left, who asks for three bread rolls in Luxembourgish, is presented as unhappy when the (happy) salesperson behind the counter uses the French shibboleth figure.

The last example (Figure 3) shows that the perception experiment, at least in part, relies on the individual ability to draw. In this case, the drawing captures an everyday interaction that is common in Luxembourg: brokering the language of interaction at the outset of a conversation. In the first scene, we see two persons negotiating their linguistic repertoires. In practice, this is often done by exchanging greetings as indicators of a person's language preference. Here, as the participant explains in the interview, the person on the left offers Luxembourgish, the other one offers English instead, and then they agree on French as a shared language. In the second scene, however, we see that when the first person speaks French, the other cannot answer properly. According to the interview, there are three layers to this reaction: a) the lack of linguistic competence (x over the mouth), b) a negative attitude towards French (thought bubble with thumb down) and, as a result, c) a lack of linguistic confidence (shrunk brain in the skull on the right).

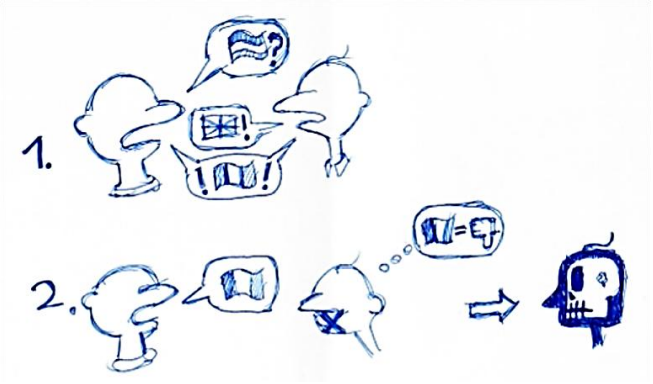

Figure 3: Participant drawing: Negotiating language preferences at the outset of a conversation.

This drawing captures significant parts of the complicated position of French in Luxembourg's multilingualism. Many Luxembourgers, especially in the younger age groups, have increasingly negative attitudes towards French, combined with a negative assessment of their (often in fact good) competence, which leads to an avoidance of the language in everyday practice. However, the fact that it is often impossible to avoid speaking French leads directly back to the discourse figures surrounding the role of French discussed above. On a side note, the drawing also works as a visual representation of the ABC model of attitudes (Eagly & Chaiken 1993).

With regard to the research question, the study confirms the central position of the discussed discourse figures in the language debate even in the outsider perspective. At the same time, the chosen method reveals differences in the motivations behind the use of certain figures, for example, in the participant's commentary on the drawing <shopping at the bakery>. In general, the outsider perspectives seem to have a more positive view on the functioning of multilingualism in the country. Since outsider perspectives are underrepresented in the Luxembourg language debate, this approach not only adds an important piece to the overall analysis, it also allows us to focus on individual stances toward the debate.

6 Discussion

The results of the individual studies demonstrate the usefulness of a multi-method approach for the examination of the public discourse on multilingualism in Luxembourg. With this in mind, I conclude by discussing key aspects of my case study, research design and analytical lens.

6.1 The role of discourse figures in the language debate

The three studies have shown that the same discourse figures can be found in different speaker groups, each with different functions, roles and positions in the discourse. On the one hand, this suggests that the selected datasets and methods can be meaningfully combined and mutually validated. Beyond that, however, this finding is revealing with regard to the structure and dynamics of the language debate. The analysis of the media corpus has shown that the main figure types articulate stances toward languages, speakers or salient features in discourse and are often complex in their semantic coding. This is the case, for example, with the figure *eis Sprooch*, in which a conception of social belonging based on legitimate linguistic competence is closely linked to a conception of language as a central element of national identity. Furthermore, the comparison of official and private actors provides insights into the uptake and circulation of ideological motifs in the discourse. To complete the picture, however, a follow-up analysis needs to investigate the temporal relationships between social contexts and groups of actors with regard to the coinage and transmission of certain figures. This would be particularly interesting for the prototypical situations to understand how they could assume such a central function in the debate. Another extension could be a comparison with the last major language debate in Luxembourg between 1974 and 1984 (Fehlen 2016), which was also preceded by a political discussion about the political participation of foreign residents, and subsequently resulted in Luxembourgish being enshrined as a language in the law. The comparison of the two debates also indicates that the public discourse on multilingualism has hardly developed structurally since then.

6.2 The concept of discourse figure

With the concept of discourse figure, the study takes a theoretical idea as a starting point for discourse analysis. In view of the practical inseparability of linguistic patterns, the ideological motifs conveyed by them and the actors responsible for them, the concept seems to make sense both theoretically and methodologically, since it enables me to describe the stances taken in discourse from a holistic perspective and to examine them in terms of the socio-pragmatic interrelationships between their components. At the same time, the use of this term might run into problems when analytically distinguishing between those components. I therefore use the term discourse figure as a label for the holistic structure and classify its components as patterns, motifs, and actors – and beyond that, topics and relations (see the definition above). Looking at the pragmatic complexity of stance taking, further research will need to shed light on the

interplay of the different factors contributing to the emergence and practical functions of figures in discourse.

6.3 The benefits and limitations of a multi-method approach

The methodological design of the study enables a differentiated examination of the socio-cultural complexity of a given discourse. The selected datasets imply a "methodological order" (Dingler 1987) for the analysis. For example, the computational modeling of discursive contexts must be preceded by the analysis of a media corpus to identify the keywords to be used for the search queries (Gabrielatos 2007). At the same time, the validity of the individual studies is limited by the fact that the same methods could not be applied to all groups of actors. The main reason for this is the specific structure of the Luxembourg public sphere, which is dominated by insiders with official functions. A promising addition to the RTL corpus could be user comments from the free newspaper *L'Essentiel*, that is published in German and French and is read mainly by cross-border commuters and foreign residents who comment on the newspaper's digital platform – mostly in French and German. In the same way, the perception study could be replicated for actors with active roles in discourse, that is, autochthonous Luxembourgers. I must also consider the fact that official and private actors in the discourse likely represent those who have either the position or a purpose to engage in the debate. For example, the language activist group Nee 2015/Wee 2050 and its constituents were the central private group of actors, despite their small size, and because of their populist agenda, they set the tone for the entire debate, with their statements being reflected in those of politicians and the media. In fact, some of the main players in this group by now represent the right-wing ADR party in different political bodies. This pattern reflects a dynamic typical of other polarized discourses (Kumkar 2025), in which a vocal minority captures all the public attention through loud statements and thus contributes to the false image of a divided society, e.g., regarding issues like climate change (Mau et al. 2023). The measurable impact of the debate on political development programs for Luxembourgish shows that such a discourse climate can nevertheless have concrete consequences.

Figure 4: Multilingual advertisement from the Lingscape repository (ID 10335); https://lingscape-app.uni.lu

7 Outlook

Based on the results of the present study, the multi-method analysis of the Luxembourg language debate could be profitably extended, for example by taking stock of the Luxembourg linguistic landscape. The *Lingscape* project provides the necessary data for such an analysis (Purschke 2017). Although the project's image database contains hardly any official statements such as election posters or stickers for the study period, the example in Figure 4 illustrates the methodological potential of this extension.

The image captures a bus stop advertisement of the German hardware store chain *Hornbach*. Using French as the matrix language of the advertisement, the company provides a variation of the language-related

type of discourse figure by giving the location of the hardware store in all three official languages: *Bertrange* (French), *Bartreng* (Luxembourgish) and *Bartringen* (German), with the addition 'we assist them all'. In doing so, the advertisement establishes an integrative image of multilingualism with Luxembourgish as a natural part – which contrasts the ideological juxtaposition of Luxembourgish and multilingualism analyzed above. At the same time, the use of French as the main language of communication suggests that its central position in the language regime is likely to continue to be a source of language-ideological debates in Luxembourg.

Author Statement

I am responsible for all aspects of this research.

Data availability

The data used in this study is available under the following conditions: The media corpus and perceptual mapping study can be made available by the author upon request. The embedding model used in the computational study is available via Zenodo using the following DOI: 10.5281/zenodo.3978066.

Endnotes

1 The documentation of the public petitions can be found online via https://www.petitiounen.lu [last accessed: 31.05.2025].

2 I did not convert the source data into lower case letters before training the model. The main reason for this lies in the fact that, due to the lack of formal orthography training, individual writing performance can be indicative of a social positioning. That is, abiding by the rules of orthography can be read as a stance toward the linguistic development and symbolic recognition of Luxembourgish.

References

Ancarmo, Clyde. 2020. Corpus-Assisted Discourse Studies. In Anna De Fina & Alexandra Georgakopoulou (eds.). *The Cambridge Handbook of Discourse Studies*, 165–185. Cambridge: Cambridge University Press.

Androutsopoulos, Jannis. 2024. Scaling as method: A three-stage, mixed-methods approach to digital discourse analysis. *Discourse, Context & Media* 62. 100817. https://doi.org/10.1016/j.dcm.2024.100817

Baker, Paul & Erez Levon. 2015. Picking the right cherries? A comparison of corpus-based and qualitative analyses of news articles about masculinity. *Discourse & Communication* 9(2). 221–236. https://doi.org/10.1177/1750481314568542

Batchelor, Jordan. 2024. Getting into bed with embeddings? A comparison of collocations and word embeddings for corpus-assisted discourse analysis. *Applied Corpus Linguistics* 4(3). 100117. https://doi.org/10.1016/j.acorp.2024.100117

Bellamy, John & Kristine Horner. 2018. Ein Mischmasch aus Deutsch und Französisch: Ideological tensions in young people's discursive constructions of Luxembourgish. *Sociolinguistic Studies* 12. 323–342.

Blommaert, Jan (ed.). 1999. *Language Ideological Debates*. Berlin: De Gruyter Mouton.

Blommaert, Jan. 2011. The long language-ideological debate in Belgium. *Journal of Multicultural Discourses* 6. 241–256. https://doi.org/10.1080/17447143.2011.595492

Borgstede, Matthias & Marcel Scholz. 2021. Quantitative and Qualitative Approaches to Generalization and Replication: A Representationalist View. *Frontiers in Psychology* 12. 605191. https://doi.org/10.3389/fpsyg.2021.605191

Bourdieu, Pierre. 1982. *Ce que parler veut dire: l'économie des échanges linguistiques*. Paris: Fayard.

Bubenhofer, Noah. 2022. Exploration semantischer Räume im Corona-Diskurs. In Heidrun Kämper & Albrecht Plewnia (eds.). *Sprache in Politik und Gesellschaft: Perspektiven und Zugänge*, 197–216. Berlin: De Gruyter. https://doi.org/10.1515/9783110774306-013

Conrad, François. 2017. *Variation durch Sprachkontakt*. Frankfurt am Main: Peter Lang.

Dingler, Hugo. 1987. Methodik statt Erkenntnistheorie und Wissenschaftslehre. In Ulrich Weiss (ed.). *Hugo Dingler. Aufsätze zur Methodik*, 1–60. Hamburg: Meiner.

Eagly, Alice & Shelly Chaiken. 1993. *The Psychology of Attitudes*. Fort Worth: Harcourt Brace & Company.

Erhart, Sabine & Fernand Fehlen. 2011. Luxembourgish: A success story? A small national language in a multilingual country. In Joshua A. Fishman & Ofelia Garcia (eds.). *Handbook of Language and Ethnic Identity*, 285–298. Oxford: Oxford University Press.

Fehlen, Fernand. 2016. Nationalsprache und nationale Identität. Die Debatten im Vorfeld des Sprachengesetzes (1974–1984). In Norbert Franz (ed.). *Identitätsbildung und Partizipation im 19. und 20. Jahrhundert. Luxemburg im europäischen Kontext*, 411–433. Frankfurt am Main: Peter Lang.

Fehlen, Fernand, Peter Gilles, Louis Chauvel, Isabelle Pigeron-Piroth, Yann Ferro & Etienne Le Bihan. 2023. *Linguistic Diversity on the Rise. RP 1ST RESULTS 2021 N°08*. Luxembourg: STATEC.

Gabrielatos, Costas. 2007. Selecting query terms to build a specialised corpus from a restricted-access database. *ICAME Journal* 31. 5–43.

Garcia, Nuria. 2014. The paradox of contemporary linguistic nationalism: the case of Luxembourg. *Nations and Nationalism* 20(1). 113–132. https://doi.org/10.1111/nana.12043

Garcia, Nuria. 2018. Multilingualism without Multiculturalism? The Case of Luxembourg. In Gillian Lane-Mercier, Denise Merkle & Jane Koutsas (eds.). *Minority Languages, National Languages, and Official Language Policies*, 63–82. Montreal: McGill-Queen's University Press.

Garg, Nikhil, Londa Schiebinger, Dan Jurafsky & James Zoue. 2018. Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Science U.S.A.* 115. E3635–E3644. https://doi.org.10.1073/pnas.1720347115

Gilles, Peter. 2015. From status to corpus: Codification and implementation of spelling norms in Luxembourgish. In Winifred Davies & Evelyn Ziegler (eds.). *Language Planning and Microlinguistics. From Policy to Interaction and Vice Versa*, 128–149. London: Palgrave Macmillan.

Gilles, Peter. 2023. Luxembourgish. In Sebastian Kürschner & Antje Dammel (eds.). *The Oxford Encyclopedia of Germanic Linguistics*. Oxford: Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.943

Grieve, Jack, Dirk Hovy, David Jurgens, Tyler S. Kendall, Dong Nguyen, James N. Stanford & Meghan Sumner (eds.). 2023. *Computational sociolinguistics*. Lausanne: Frontiers Media SA. https://doi.org/10.3389/978-2-8325-1760-4

Gumperz, John. 1982. *Discourse Strategies*. Cambridge: Cambridge University Press.

Hodge, Gabrielle & Goico, Sara A. 2022. Natural and elicited: Sign language corpus linguistics and linguistic ethnography as complementary methodologies. *Journal of Sociolinguistics* 26. 126–136. https://doi.org/ 10.1111/josl.12523

Horner, Kristine & Jean-Jacques Weber. 2008. The language situation in Luxembourg. *Current Issues in Language Planning* 9. 69–128.

Jaffe, Alexandra. 2009. *Stance: Sociolinguistic perspectives*. Oxford: Oxford University Press.

Keller, Reiner. 2011. *Wissenssoziologische Diskursanalyse. Grundlegung eines Forschungsprogramms*. 3rd edition. Wiesbaden: VS Springer.

Kircher Ruth & James Hawkey. 2022. Mixed-Methods Approaches to the Study of Language Attitudes. In Ruth Kircher & Lena Zipp (eds.). *Research Methods in Language Attitudes*, 330–345. Cambridge: Cambridge University Press.

Kozlowski, Austin. C., Matt Taddy & James A. Evans. 2019. The geometry of culture: analyzing the meanings of class through word embeddings. *American Sociological Review* 84. 905–949.

Kress, Gunther & Theo van Leeuwen. 2001. *Multimodal discourse*. London: Bloomsbury Academic.

Kumkar, Nils. 2025. *Polarisierung. Die Ordnung der Politik*. Berlin: Edition Suhrkamp.

Lutgen, Anne-Marie, Alistair Plum, Christoph Purschke & Barbara Plank. 2025. Neural Text Normalization for Luxembourgish Using Real-Life Variation Data. *Proceedings of the 12th Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial)*, Abu Dhabi, 115–127.

Mau, Steffen, Thomas Lux & Steffen Westheuser. 2023. *Triggerpunkte. Konsens und Konflikt in der Gegenwartsgesellschaft*. Berlin: Edition Suhrkamp.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg Corrado & Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. *Proceedings of the 26th International Conference on Neural Information Processing Systems*, Volume 2, 3111–3119. https://0.5555/2999792.2999959

Naglo, Kristian. 2007. *Rollen von Sprache in Identitätsbildungsprozessen multilingualer Gesellschaften in Europa. Eine vergleichende Betrachtung Luxemburgs, Südtirols und des Baskenlands*. Frankfurt am Main: Peter Lang.

Partington, Alan, Alison Duguid & Charlotte Taylor. 2013. *Patterns and Meanings in Discourse: Theory and Practice in Corpus-Assisted Discourse Studies (CADS)*. Amsterdam: John Benjamins.

Péporté, Pit, Sonja Kmec, Benoît Majerus & Michel Margue. 2010. *Inventing Luxembourg. Representations of the Past, Space and Language from the Nineteenth to the Twenty-First Century*. Boston: Brill.

Philips, Susan. 2020. Anthropological Discourse Analysis and the Social Ordering of Gender Ideology. In Judith Baxter & Jo Angouri (eds.). *The Routledge Handbook of Language, Gender, and Sexuality*, 136–149. London: Routledge.

Purschke, Christoph. 2014. Wort und Totschlag – Zur sozio-symbolischen Bedeutung sprachlicher Divergenz. In Thomas Krefeld & Elissa Pustka (eds.). *Perzeptive Linguistik,* 51–64. Stuttgart: Steiner.

Purschke, Christoph. 2017. Crowdsourcing the linguistic landscape of a multilingual country. Introducing Lingscape in Luxembourg. *Linguistik Online* 85(6). 181–202.

Purschke, Christoph. 2020a. Fescher als dein Schatten. Zur Präsenz des Deutschen in Österreich in der Alltagspraxis. In Markus Hundt, Andrea Kleene, Albrecht Plewnia & Verena Sauer (eds.). *Regiolekte. Objektive Sprachdaten und subjektive Wahrnehmung*, 315–343. Tübingen: Narr Francke Attempto.

Purschke, Christoph. 2020b. Attitudes towards multilingualism in Luxembourg. A comparative analysis of online news comments and crowdsourced questionnaire data. *Frontiers in Artificial Intelligence* 3. 536086. https://doi.org/10.3389/frai.2020.536086

Purschke, Christoph. 2023. Diskurs-Figuren. Wie Politik und Öffentlichkeit über Sprache sprechen*. Hémecht. Zeitschrift für Luxemburger Geschichte* 3/2023. 313–330.

Purschke, Christoph & Peter Gilles. 2023. 46. Sociolinguistics in Luxembourg. In Martin J. Ball, Rajend Mesthrie & Chiara Meluzzi (eds.). *The Routledge Handbook of Sociolinguistics around the World*, 2nd edition, 542–549. London: Routledge.

Purschke, Christoph & Mirjam Schmalz. 2022. Mapping knowledge and perceptions of language: societal multilingualism and its socio-pragmatic grounding. In Andreas Jucker & Heiko Hausendorf (eds.). *Pragmatics of Space*, 679–714. Berlin: De Gruyter. https://doi.org/10.1515/9783110693713-021

Rivera Cosme, Gabriel. 2023. *Language Policy in Luxembourg and the German-speaking Community of Belgium: Ideologies of Language*. University of Luxembourg. (PhD dissertation).

Schwarz, Ori. 2021. *Sociological Theory for Digital Society: The Codes that Bind Us Together.* Hoboken: Wiley.

Spitzmüller, Jürgen. 2023. Metapragmatische Positionierung. Reflexive Verortung zwischen Interaktion und Ideologie. In Mark Dang-Anh (ed.). *Politisches Positionieren. Sprachliche und soziale Praktiken*, 39–66. Heidelberg: Winter.

Spitzmüller, Jürgen & Ingo Warnke. 2011. *Diskurslinguistik. Eine Einführung in Theorien und Methoden der transtextuellen Sprachanalyse*. Berlin: De Gruyter.

STATEC. 2023. *Luxembourg in Figures 2023*. Luxemburg: STATEC.

Tannen, Deborah, Heidi Hamilton & Deborah Schiffrin (eds.). 2015. *The Handbook of Discourse Analysis*. Hoboken: Wiley.

Tusting, Karin (ed.). 2020. *The Routledge handbook of linguistic ethnography*. London: Routledge.

Vessey, Rachelle. 2017. Corpus Approaches to Language Ideology. *Applied Linguistics* 38(3). 277–296. https://doi.org/10.1093/applin/amv023

Watts, Richard. 1999. The ideology of dialect in Switzerland. In Jan Blommaert (ed.). *Language Ideological Debates*, 67–104. Berlin: De Gruyter Mouton.

Wertheimer, Max. 1925. *Über Gestalttheorie, Vortrag vor der Kant-Gesellschaft, Berlin am 17. Dezember 1924*. Erlangen: Verlag der Philosophischen Akademie.