

“BABEȘ-BOLYAI” UNIVERSITY, CLUJ-NAPOCA  
FACULTY OF LETTERS  
DOCTORAL SCHOOL OF LINGUISTIC AND LITERARY STUDIES

PhD. Thesis in the Field of Philology

**ENGLISH DISCOURSE MARKERS IN COLLOQUIAL ROMANIAN.  
A CASE STUDY ON TIKTOK**  
*Thesis Summary*

Scientific advisor:

Prof. Univ. Dr. Ștefan OLTEAN

Doctoral student:

Georgiana-Anamaria SĂLĂGEAN (MILEA)

Cluj-Napoca

2026

## TABLE OF CONTENTS

|  |    |
|--|----|
| <b>Introduction</b> .....  | 5  |
| <b>Methodology</b> .....   | 11 |
| <b>Chapter 1 – Theoretical Background. Concepts and Perspectives</b> ..... | 15 |
| 1.1. Functions of Language .....   | 15 |
| 1.1.1. Karl Bühler .....   | 16 |
| 1.1.2. Roman Jakobson .....  | 17 |
| 1.1.3. Social Functions .....  | 20 |
| 1.1.4. Metafunctions of Language .....                                     | 22 |
| 1.2. Discourse and Discourse Analysis .....                                | 23 |
| 1.3. English as a Contact Language .....                                   | 25 |
| 1.4. English Influence on Romanian and ‘Glocal’ Identity .....             | 26 |
| 1.5. Borrowings and Code-switching .....                                   | 31 |
| 1.6. Discourse Markers .....   | 34 |
| 1.7. Colloquial Language .....   | 39 |
| 1.8. TikTok Insights .....   | 44 |
| 1.9. A Sociolinguistic Perspective and Speech Communities .....            | 53 |
| 1.10. Superdiversity and Social Media .....                                | 55 |
| 1.11. Social Media and Pragmatics .....                                    | 56 |
| 1.12. Plurilingualism and Translanguaging .....                            | 57 |
| 1.13. Subcultural Phenomenon .....   | 58 |
| 1.14. Language and Identity .....  | 59 |

|  |            |
|--|------------|
| 1.15. Partial Conclusions .....  | 60         |
| <b>Chapter 2 – The Corpus. A Foundation for Analysis .....</b>                           | <b>64</b>  |
| 2.1. Transcription Conventions .....   | 76         |
| 2.2. Classification of Discourse Markers .....   | 78         |
| 2.3. Transcriptions .....  | 79         |
| <b>Chapter 3 – Exploring Data. Insights from Corpus and Questionnaire Analysis .....</b> | <b>84</b>  |
| 3.1. Classification of Transcriptions .....  | 84         |
| 3.2. Defining the Video Clip Types .....   | 87         |
| 3.3. Initial Insights and Trends Observed from the Corpus Data .....                     | 88         |
| 3.4. Limitations of the Corpus .....   | 91         |
| 3.5. Addressing Limitations in Future Research .....                                     | 94         |
| 3.6. Analysis .....  | 96         |
| 3.7. Conclusions .....   | 124        |
| 3.8. The Questionnaire .....   | 129        |
| <b>Chapter 4 – Findings and Discussion .....</b>   | <b>154</b> |
| 4.1. Key Findings .....  | 154        |
| 4.2. Theoretical Implications .....  | 160        |
| 4.3. Cultural and Identity Implications .....  | 163        |
| 4.4. Social Dynamics and Code-Switching .....  | 168        |
| 4.5. Practical Communication and Linguistic Economy .....                                | 170        |
| 4.6. Educational and Policy Implications .....   | 172        |
| 4.7. Directions for Future Research .....  | 174        |
| 4.8. Conclusion .....  | 176        |

|                                  |     |
|----------------------------------|-----|
| <b>General Conclusions</b> ..... | 179 |
| <b>Bibliography</b> .....        | 186 |
| <b>Appendix</b> .....            | 192 |

**Keywords:** English discourse markers, digitally mediated communication, TikTok discourse, code-mixing, discourse-pragmatic functions, colloquial Romanian, online communities, sociolinguistics of social media.

The thesis *English Discourse Markers in Colloquial Romanian: A Case Study on TikTok* motivates an investigation into how English discourse markers have made their way into colloquial Romanian used on digital platforms, with TikTok as a focal site of analysis. It situates ongoing language change within sociolinguistic and cultural dynamics shaped by globalisation and digitally mediated communication, arguing that the colloquial register is particularly conducive to contact-induced innovation, including lexical borrowing, code-mixing, and pragmatic transfer. In digitally mediated settings, English can function as a high-prestige contact resource circulating through media and technology (Crystal 2003; Vennemann 2011). Rather than focusing on speakers' individual linguistic profiles, the study concentrates on observable language-contact practices and their interactional functions in TikTok discourse.

Building on prior work — namely, the bachelor's thesis “Current Trends of Colloquial Language” (Sălăgean 2019) and master's dissertation “Linguistic Expression and Translation Loans from English” (Sălăgean 2021) — the thesis identifies a research gap concerning the pragmatic use of English markers in contemporary Romanian online interaction. Its central aim is to analyse Romanian TikTok creators' deployment of markers such as *so*, *you know*, and *anyway* using a purpose-built corpus of transcribed video clips, combining quantitative frequency occurrence analyses with qualitative, context-sensitive analysis. To complement creator-focused data, a questionnaire targeting non-creators explores how factors such as age, gender, education, and social media exposure relate to self-reported code-mixing tendencies, indicating the limits of assessing the creators' biographies from public content.

The argument emphasises that digitally mediated discourse often blends features associated with spoken and written communication (Crystal 2004), favouring efficiency, stance-marking, and audience alignment. In this environment, English discourse markers can contribute to discourse organisation, interpersonal positioning, and community-oriented styles, sometimes competing with Romanian functional equivalents. This can also be read through the lens of ‘glocal’ style-making, where local repertoires are reworked through globally circulating semiotic resources (Robertson 1992; Roudometof 2016). The study's roadmap previews the theoretical framing, corpus design, transcription procedures, and analytic steps, and a discussion of implications for digital pragmatics, language contact, identity work, and the sociolinguistics of social media.

Against this background, the next section outlines the methodological design adopted to operationalise these aims and to analyse English discourse markers in Romanian TikTok discourse.

## **Methodology**

The thesis adopts a mixed-methods design combining corpus-based analysis with observational and demographic components to examine English discourse markers in colloquial Romanian TikTok content. TikTok is treated as a multimodal environment that encourages short-form, audience-oriented performances and routinised interactional formats, which makes it suitable for pragmatic analysis (Schellewald 2021; Crystal 2004). The corpus was compiled from publicly available video clips produced by Romanian creators aged 14–30. This age range was used as a sampling criterion because TikTok use and routine exposure to English-mediated content are particularly salient in this demographic; age information was considered only where it was publicly available (e.g., self-disclosed in profiles or other accessible public sources) and was treated as an approximate selection parameter rather than an independently verified variable. Private accounts were excluded for ethical reasons, and the study avoids assigning individual-level labels (e.g., bilingual or plurilingual) because the corpus does not include linguistic biographies or independent proficiency measures. The analysis, therefore, targets observable language-contact practices, including code-mixing, code-switching, and the insertion of markers such as *Wow*, *Oh my God*, *so*, *okay*, and *like*, capturing colloquial spoken features of digitally mediated discourse.

Sampling aimed to include a range of communicative genres (informal announcements, advertising, personal vlogs, and opinion narratives) to document authentic uses of English markers embedded in Romanian discourse. Quantitative procedures were used to identify marker frequency and distribution across the corpus, while qualitative pragmatic analysis examined their contextual functions (e.g., discourse organisation, emphasis, stance, and audience engagement) in Romanian-dominant interaction. This approach follows work in social-media pragmatics that prioritises context, participation frameworks, and platform-specific cues in interpretation (Bărbuleț 2013; Chung and Liu 2023). Ethical safeguards underpin the design: only public data were analysed, and confidentiality is ensured for any questionnaire-based component. To complement creator-focused data, a questionnaire targeting non-creators explores how demographic factors (age, gender, education, and social media exposure) relate to self-reported language-contact behaviour.

I aimed to keep the analysis firmly grounded in the empirical data: the selection of clips, the transcription procedures, and the coding of discourse markers are described in sufficient detail for the study to be replicated on a larger corpus. The methodology links language-contact theory to corpus evidence and sets out steps that can be replicated (sampling, transcription, coding) for investigating how digital environments facilitate multilingual discourse practices and identity work in contemporary Romanian.

## **Chapter 1 – Theoretical Background. Concepts and Perspectives**

Chapter 1 sets out the theoretical background used to analyse English discourse markers in Romanian TikTok discourse. It draws mainly on pragmatics, sociolinguistics, and discourse analysis, because these strands offer the most direct tools for explaining how markers work in interaction and how they circulate in language-contact settings online. The aim is not to *diagnose* speakers' linguistic profiles, but to clarify the concepts and categories that will later be applied to the corpus.

The chapter opens with functional accounts of language. Bühler's Organon Model is introduced to highlight the fact that utterances typically do more than convey information: they can express the speaker's stance, represent states of affairs, and address an audience to prompt reactions. Jakobson's model extends this view by linking functions to key elements of communication (addresser, addressee, message, context, code, and contact). For the present topic, the phatic and interpersonal dimensions are often especially relevant, since TikTok creators often speak to an audience, yet they still design their speech to sustain attention, invite responses, and build rapport.

The sociolinguistic discussion develops this further by focusing on language as a social practice. Rather than treating meaning as purely lexical, here I emphasise how linguistic choices can signal relationships, positioning, and group membership (Barker 1945; Hudson 1996). Halliday's metafunctions then provide a bridge between form and function: the ideational metafunction relates to how experience is represented; the interpersonal metafunction to how speakers align with, address, or evaluate; and the textual metafunction to how discourse is organised and made coherent. This perspective is useful for discourse markers because many of

them are small items that nevertheless contribute to cohesion and to the management of interpersonal meaning. This focus on function and use also aligns with Romanian perspectives on everyday language, understood as a context-sensitive variety shaped by routine interaction and communicative purpose (Oltean 1977). This conceptual bridge helps motivate the thesis's emphasis on colloquial registers as the primary site where pragmatic resources such as discourse markers become visible and recurrent.

A separate section clarifies the notion of discourse and the rationale for using discourse analysis. Following Schiffrin, discourse is treated as structured, meaning-bearing, and action-oriented; it is not reducible to isolated sentences. Brown and Yule's emphasis on language in use supports the chapter's broader claim that context is indispensable when interpreting pragmatic items. This point becomes even more important in digital environments. This aligns with discourse-analytic approaches that treat meaning as inseparable from situated use (Brown and Yule 2020 [1983]). Although TikTok is a platform, the data analysed in this thesis are spoken performances in video clips. They often combine spoken and written cues (speech, captions, on-screen text), which resonates with Crystal's observation that online communication can display a *spoken* dimension even when it is technologically mediated (Crystal 2004).

The section then turns to language contact and to English as a global resource. English is discussed as a contact language whose influence can spread through media, technology, and cultural prestige (Crystal 2003; Vennemann 2011; Carter 1993). The Romanian setting is contextualised through research on anglicisms and pragmatic borrowings, including discourse-marker-focused studies (Avram 1997; Cojocaru 2020; Niculescu-Gorpin and Vasileanu 2020). Here, I distinguish borrowing, code-mixing, and code-switching, using the labels in a strictly descriptive way. The emphasis is methodological: the analysis focuses on what can be observed in the corpus (insertions, alternations, and marker use in Romanian-dominant talk), without drawing claims about individual competence or stable speaker typologies.

Discourse markers are defined as pragmatic items with primarily procedural meaning, and the chapter reviews core typologies that are later used to interpret the corpus. In the thesis, markers are treated mainly as procedural cues that guide interpretation rather than adding propositional content (Fraser 1996; Schiffrin 1987). A practical point is stressed here: markers help speakers connect segments of talk, signal transitions, manage emphasis, and shape audience-directed

interaction. They are therefore particularly informative in short, fast-paced TikTok genres where creators need to keep speech flowing and maintain engagement.

### **How the Framework Is Used in the Empirical Chapters**

The categories introduced in Chapter 1 are not treated as abstract labels, but as tools for annotation and interpretation in Chapters 2–4. In the corpus analysis, each English marker is examined in context with attention to (a) discourse organisation (e.g., sequencing, topic shifts, elaboration, inferencing), (b) stance and interpersonal work (e.g., evaluation, mitigation, intensification, alignment), and (c) interaction management (e.g., attention-getting, floor-holding, audience address). These functional observations are combined with frequency and distributional patterns across clip types, which allows the study to connect quantitative tendencies with qualitative pragmatic effects.

The final sections bring in concepts that are particularly relevant for social media environments: speech communities, superdiversity, pragmatics, translanguaging, subcultures, and identity (Gumperz 1968; Hymes 1974; Blommaert 2011; Bucholtz and Hall 2004). The chapter concludes by arguing that an integrated framework is needed to capture what discourse markers do in Romanian TikTok discourse and why English items may be selected as part of a broader repertoire of online language-contact practices.

On this basis, Chapter 2 describes how the TikTok corpus was built and standardised so that these categories can be applied consistently to the data.

## **Chapter 2 – The Corpus. A Foundation for Analysis**

Chapter 2 describes how the corpus was assembled and prepared for analysis. It introduces the TikTok corpus, clarifies the selection principles, and sets out the transcription and coding procedures used to ensure comparability across clips and genres.

### **Corpus Design and Selection Principles**

The corpus is built from publicly available TikTok video clips produced by Romanian content creators within the 14–30 age range, selected when Romanian remains the dominant code, and English appears primarily as insertions within Romanian talk (including discourse markers).

Private accounts and restricted content were excluded, and the corpus relies only on material accessible without special permissions. This sampling logic keeps the corpus ethically manageable and aligns with the methodological stance of the thesis: the analysis targets observable language-contact practices rather than biographical claims about speakers.

In terms of communicative settings, the corpus was designed to include recurring TikTok formats in which informal speech is prominent and where discourse markers are likely to occur naturally: personal vlogs (including GRWM), tutorials or routines, hauls and unboxings, and short commentary or promotional sequences. These genres are relevant because they regularly involve step-by-step narration, rapid topic movement, evaluative peaks, and direct audience address—features that create strong interactional environments for marker use.

### **Transcription and Annotation Conventions**

A central part of the chapter presents the transcription conventions used throughout the study. The transcription system prioritises readability but captures features that are important for pragmatic interpretation, such as pauses, emphatic stress, lengthening, laughter, and intonation movement when clearly audible. Unclear stretches are marked transparently, and occasional omissions are signalled when they do not affect the interpretation of discourse-marker use. In the running text, discourse markers are presented in *italics* (e.g., *so, you know, anyway*) to signal that they are cited as linguistic items rather than used as propositional content.

English discourse markers are consistently highlighted in the transcripts to support later comparisons across clips and contexts. The goal is not in-depth phonetic detail, but a stable representation of interactional cues that allow markers to be interpreted in relation to sequencing, evaluation, and audience-oriented talk.

### **Marker Set and Classification Logic**

Chapter 2 also clarifies the analytic focus of the corpus by defining the set of 15 English discourse markers tracked in the study. I distinguish discourse markers from lexical borrowings using three practical criteria: optionality, limited propositional content, and recurrent discourse positions (e.g., transitions, evaluations, audience-facing sequences). The distinction is informed by Romanian research on how English items may undergo pragmaticisation in Romanian usage (Cojocaru 2020; Niculescu-Gorpin and Vasileanu 2020).

For the empirical chapters, markers are organised functionally as an analytic aid rather than as a rigid taxonomy. In practice, the classification supports consistent coding of how markers contribute to discourse organisation (e.g., sequencing, topic shifts, elaboration), stance and interpersonal meaning (e.g., intensification, mitigation, alignment), and interaction management (e.g., attention-getting, audience address, floor-holding). This functional organisation creates a direct bridge between the theoretical concepts established in Chapter 1 and the procedures applied to the corpus.

### **Corpus Presentation and the Role of Sample Transcriptions**

The section closes by presenting the ten transcriptions that form the core corpus for the study. These examples serve a dual role: they document the corpus transparently and illustrate how the transcription and coding conventions operate in real spoken material. Across the samples, English discourse markers tend to be especially visible at points of transition, strong evaluation, and direct address to viewers. This provides a clear empirical foundation for the analytic work in Chapters 3 and 4.

### **Ethics and Transparency**

Finally, Chapter 2 reiterates the ethical parameters of the research design. Only public video content is used, identifying details are minimised where appropriate, and the study avoids assigning individual-level labels (e.g., bilingual or plurilingual) beyond what can be observed in the clips. Overall, the chapter lays out procedures that make the corpus and subsequent analyses traceable, consistent, and replicable.

With the corpus and coding procedures in place, Chapter 3 turns to distributional patterns and contextual uses of the markers across the ten transcriptions, supplemented by questionnaire data.

## **Chapter 3 – Exploring Data. Insights from Corpus and Questionnaire Analysis**

Chapter 3 presents the empirical analysis of the study, drawing on both the transcribed TikTok corpus and the questionnaire administered to non-creators. It maps where exactly English discourse markers occur in Romanian-dominant TikTok talk and what kinds of interactional work

they do in context. It combines quantitative summaries with qualitative, context-based interpretation, so that distributional patterns can be related to the communicative routines of different clip types and to audience-oriented speaking on TikTok. For terminological clarity, code-mixing and code-switching are discussed with reference to established typologies of code-switching (Poplack 1980).

Percentages reported in this chapter are based on token counts in the transcriptions. For each transcript, I counted the total number of words and then calculated the share of Romanian tokens, English tokens, and English discourse markers within the spoken material. These figures are used descriptively to compare clips in this corpus.

### **Transcription Classification and Video-Genre Coding**

The analysis begins by classifying the ten transcriptions according to video type and communicative purpose. Each transcript is coded for its main genre (e.g., makeup routine, tutorial, GRWM, haul, unboxing, advertising, promotional content, and rating or evaluation) and for the English discourse markers that occur in the spoken sequence. This step provides an organisational baseline for later comparisons and supports the identification of recurring environments for marker use, such as step-by-step explanations, evaluative sequences, and direct address to viewers.

### **Trends and Distributional Patterns in the Corpus**

Across the corpus, the style is informal, fast-paced, and strongly audience-oriented. English occurs alongside Romanian primarily as insertions in Romanian-dominant talk. In the ten transcribed clips, English discourse markers are particularly visible at points where speakers move from one subtopic to another, shift from description to evaluation, or turn explicitly to the audience. A small set of markers recurs across multiple clips, suggesting that certain English items are routinely available as pragmatic resources in this type of TikTok speech. Descriptive counts show that the share of English tokens varies by clip and by genre segment rather than being evenly distributed; higher concentrations tend to occur where talk becomes more evaluative or interactional, for instance, during moments of surprise, strong approval or disapproval, or rapid topic movement. These quantitative patterns are treated as corpus-internal observations rather than general claims about Romanian youth language. The clustering of markers in routinised slots is

compatible with accounts of superdiverse online repertoires (Blommaert 2011; Leppänen, Peuronen, and Westinen 2018).

### **Limitations and Future Directions**

I also address limitations that follow from this sample and the platform context. With ten transcriptions, the analysis captures recurring practices in this corpus but cannot support broad generalisations beyond it. Moreover, TikTok's visibility mechanisms and trend cycles can affect what types of content are easiest to observe and collect, potentially favouring certain genres. The chapter, therefore, outlines extensions such as expanding the number of clips, widening genre coverage, sampling across a broader set of creators, and adding cross-platform comparisons. It also notes that interviews or additional qualitative instruments could provide a clearer perspective on creators' own reported norms, without relying on biographical inference from public content alone.

### **Contextual Analysis and Interim Conclusions**

The core analytic section discusses the clips in detail and interprets marker use in context. Building on Chapter 1, markers are described in terms of their contribution to (a) discourse organisation, (b) stance and interpersonal meaning, and (c) interaction management. Function is linked to concrete positions in the discourse rather than to isolated word lists. Several recurrent patterns are highlighted: markers support discourse organisation in instructional or list-based genres (e.g., *so*, *anyway* at transitions), contribute to stance marking in evaluative peaks (e.g., *Oh my God*, *Wow* as intensifiers and alignment cues), and help manage interaction through audience address and rapport (e.g., *girls* as a conversational framing device). Collectively, these observations support the chapter's claim that English discourse markers function as practical pragmatic tools in Romanian-dominant TikTok talk, particularly in short-form genres where rhythm and engagement are central.

### **Questionnaire Results**

Chapter 3 closes with questionnaire findings (N = 70), used as a complementary perspective on awareness and attitudes. Many respondents report frequent exposure to English insertions in online contexts, and several indicate that they themselves use such items, sometimes intentionally (e.g., speed, humour, online tone) and sometimes habitually. Attitudes are mixed: some participants treat English insertions as normal in social-media discourse, while others report

annoyance when they perceive them as excessive or unnecessary. Open-ended responses add nuance by pointing to platform routines, pop-culture exposure, and perceived peer norms as relevant background factors. Thus, Chapter 3 establishes the main empirical tendencies that inform Chapter 4, linking corpus-based patterns to user perceptions of digitally mediated language-contact practices.

These corpus patterns and reported attitudes provide the empirical basis for the broader interpretation developed in Chapter 4.

## **Chapter 4 – Findings and Discussion**

Chapter 4 integrates the corpus-based analysis with the questionnaire results to explain what English discourse markers do in Romanian-dominant TikTok talk and how such language-contact practices are perceived by young users. The discussion builds on the conceptual tools developed in Chapter 1, but it stays close to what can be observed in the data: recurrent marker positions, interactional routines, genre constraints, and audience design on a platform where speech is produced for an unseen but constantly addressed public. Rather than assigning individual-level labels (e.g., bilingual or plurilingual) to creators, the chapter treats English marker use as a set of situated practices that become conventionalised in particular online genres and communicative moments. These choices can be discussed as identity work accomplished through stance, alignment, and indexical resources available in the platform repertoire (Bucholtz and Hall 2004).

### **Key Findings and Their Distribution across Genres**

Across the ten transcriptions, English discourse markers are not randomly scattered; they cluster in predictable environments shaped by short-form storytelling, rapid sequencing, and the need to hold attention. In tutorial-like formats (routines, GRWM, demonstrations), markers frequently appear at points where speakers move from one step to the next, resume a thread after a small digression, or close a segment before continuing. In hauls and unboxings, marker density tends to increase around evaluative peaks, moments of surprise, approval, disappointment, or humorous exaggeration, where the creator's stance becomes part of the entertainment value of the clip. Promotional sequences show a similar pattern, but with a stronger persuasive orientation:

markers contribute to pacing, emphasis, and the management of familiar sales-talk routines. Taken together, the corpus suggests that English discourse markers function as practical resources for organising talk, packaging evaluation, and maintaining a conversational frame in a context that rewards speed, clarity, and engagement.

### **Discourse-Pragmatic Functions Used as Analytic Instruments**

The chapter applies the functional lens introduced in Chapter 1 as a concrete analytic instrument in the empirical reading of each transcript. Marker uses are interpreted along three intersecting dimensions. First, in terms of discourse organisation, markers frequently serve as procedural cues that manage coherence: they support sequencing, signal topic shifts, and guide the listener through compressed narratives and list-like structures typical of TikTok speech. Second, in terms of stance and interpersonal work, markers often index evaluation, intensification, mitigation, or alignment; interjections and short formulas (e.g., *Oh my God*, *Wow*) regularly appear as stance boosters at moments where the creator displays affect and invites the viewer to share it. Third, in terms of interaction management, markers contribute to audience design by simulating dialogue in asynchronous communication. Address terms and phatic cues (e.g., *girls*) can position the viewer as a participant, while other markers help speakers hold the floor, keep rhythm, and avoid breakdowns in fluency during fast-paced delivery. This three-part functional grid is used consistently across the corpus so that qualitative observations (what a marker accomplishes in context) can be related to quantitative patterns (where markers recur and in which clip types).

To make this operationalisation visible, in this chapter, I discuss recurrent items as case-like illustrations of function-in-context. The marker *so* frequently appears at moments where speakers move from background explanation to the next step in a routine or demonstration, functioning as a sequencing cue that supports compressed instruction and forward movement. In these contexts, it is less a borrowed content word and more a procedural signal that frames what follows as continuation, consequence, or next action, which aligns with the platform's preference for pace and clarity. By contrast, *anyway* tends to occur where speakers re-enter a main thread after a brief aside or where they close an evaluative segment before shifting topic; it therefore operates as a boundary marker that supports coherence management and smooth topic movement in short-form talk. Markers such as *Oh my God* and *Wow* cluster around evaluative peaks, moments of surprise, strong approval or disapproval, or humorous exaggeration, where they function primarily

as stance intensifiers and help package affect in an easily readable way. Finally, address terms and phatic cues such as *girls* are recurrent in audience-facing sequences (openings, engagement checks, closing prompts), where they support interaction management by explicitly constructing a participation framework and reinforcing a community-oriented, conversational style despite the asynchronous setting. In short, these examples show how the analytic framework is used in the empirical chapters: markers are interpreted not only as English insertions, but as tools for discourse organisation, stance work, and audience design within Romanian-dominant TikTok communication.

### **Social Meanings, Platform Norms, and Identity Work**

Beyond their local pragmatic functions, I discuss how English markers may contribute to the social meaning of TikTok talk. In the corpus, English items are compatible with an online style oriented toward trend-awareness, performative informality, and affiliation with wider platform communities. The analysis does not treat these meanings as fixed or universal; instead, it shows that the same marker can participate in different identity work depending on the clip's genre, topic, and the speaker's immediate stance. Importantly, I avoid framing English markers as replacements for Romanian in a straightforward way. In many cases, English items appear as recognisable micro-resources that package transitions, stance, or rapport efficiently, while Romanian remains the matrix of the interaction. The discussion, therefore, links marker choice to the communicative ecology of the platform: compressed time, public address, routinised genres, and the circulation of globalised media templates.

### **Code-Mixing, Alternation, and the Shape of Romanian-Dominant Talk**

Chapter 4 also clarifies how the patterns observed relate to broader language-contact processes. In the material analysed here, English occurs primarily as insertions in Romanian-dominant talk, with discourse markers occupying positions that are typical of pragmatic items: clause edges, turn-launching slots, evaluative bursts, and audience-address sequences. Where alternation between Romanian and English occurs, it is treated descriptively as part of a repertoire of online speaking practices, often motivated by the communicative demands of the genre (quoting, emphasis, named formats, catchphrases) rather than as evidence for stable competence profiles. This approach keeps the discussion consistent with the methodological decision to focus on observable behaviour and context rather than on speaker typologies.

## **Questionnaire Results and the Interpretive Bridge to Corpus Data**

The questionnaire findings (N = 70) provide a complementary perspective on awareness and attitudes. Many respondents report frequent exposure to English insertions in social-media discourse and indicate that they themselves use such items, sometimes intentionally (for speed, humour, *online tone*) and sometimes habitually. At the same time, attitudes are not uniform: several respondents accept English insertions as normal within TikTok-like contexts, while others describe irritation when they perceive the practice as excessive, performative, or *forced*. Open-ended responses add nuance by pointing to peer norms, pop-culture circulation, and the communicative efficiency of short English formulas as relevant background factors. I read these responses alongside the corpus patterns: what participants describe as natural online speech corresponds to the corpus finding that markers cluster in routinised interactional slots, while negative evaluations often concern density and appropriateness of use rather than the mere presence of English.

## **Implications for Digital Pragmatics and Language-Contact Research**

Based on the combined evidence, Chapter 4 argues that English discourse markers in Romanian TikTok talk are best analysed as pragmatic tools that facilitate linguistic economy and interactional design in digitally mediated speech. The study does not claim language decline or uniform change; instead, it shows how platforms can accelerate the circulation and normalisation of particular pragmatic forms, especially those that are short, recognisable, and well-suited to fast-paced audience-oriented genres. The chapter also underlines the methodological value of combining distributional measures with close contextual analysis: frequencies help identify recurring tendencies, while pragmatic reading explains what those tendencies accomplish in concrete interaction.

## **Practical Relevance and Educational Reflections**

The practical discussion is deliberately cautious. Rather than proposing restrictive language policies for everyday online talk, the chapter suggests that register awareness is a more productive response: speakers can learn to recognise when colloquial platform routines are appropriate and when formal or academic contexts call for different pragmatic resources. In this sense, English marker use becomes a useful case study for discussing register, audience design, and style-shifting in contemporary Romanian. From an educational perspective, this connects to

additive views of multilingual development, while remaining attentive to potential register imbalance across contexts (Cummins 2017; Grosjean 2010).

### **Directions for Future Research**

Finally, Chapter 4 outlines avenues for further work: enlarging the corpus, widening genre coverage, adding cross-platform comparisons, and incorporating longitudinal sampling to observe whether marker routines stabilise or shift across trend cycles. Additional qualitative instruments (e.g., interviews) are also proposed as a way to capture creators' reported norms and motivations without relying on inference from public clips alone. These extensions would help refine the account of how digital ecologies shape language-contact practices and how such practices are evaluated by different user groups.

Based on these findings and discussion points, the thesis closes by summarising the main contributions, limitations, and directions for future research.

### **General Conclusions**

In this thesis, I examined how English discourse markers appear in Romanian-dominant TikTok speech and what they seem to accomplish in short, audience-oriented clips. Because the corpus for this study consists of publicly available videos and does not include linguistic biographies or independent proficiency measures, I did not attempt to assign individual-level labels (such as bilingual or plurilingual) to specific creators. I kept the focus on what can be observed in the recordings: English insertions, occasional alternations, and recurrent marker routines in Romanian talk.

One clear result is that these markers are not scattered at random. In the ten transcribed clips, they tend to concentrate in recognisable positions: when speakers move from one step to the next, when they shift topic, when they return to the main thread after a brief aside, and when they reach evaluative peaks (surprise, excitement, strong approval or disapproval). They also appear in moments where creators explicitly address the viewer and try to maintain a conversational tone despite the asynchronous format of TikTok.

Interpreting these uses through the framework developed in Chapter 1, I repeatedly found that English discourse markers help organise talk, package stance, and support audience-oriented interaction. These functions recur throughout the corpus, and in many cases, the English items function less like borrowed vocabulary and more like procedural signals that keep speech flowing and audience-oriented.

Methodologically, the study combines a small corpus with a complementary questionnaire. The corpus component includes ten transcriptions and tracks 15 recurrent English discourse markers. In Chapter 3, the quantitative summaries are based on token counts: for each transcription, I counted the total number of words and calculated the proportions of Romanian tokens, English tokens, and English discourse-marker tokens. These figures are used descriptively to compare clips within this corpus; they are not meant to support broad generalisations about Romanian youth as a whole. The questionnaire (N = 70, ages 14–30) adds another layer by capturing self-reported exposure, use, and attitudes.

The questionnaire responses broadly align with the corpus patterns. Many respondents describe English insertions as common in social-media registers and associate them with speed, humour, convenience, or *online tone*. At the same time, evaluations are mixed. What people criticise is usually not the presence of English itself, but the feeling that it is excessive, performative, or unnecessary in a given moment. That also fits what the corpus shows: marker density varies by clip and tends to increase when the talk becomes more evaluative or more directly oriented to the audience.

Broadly speaking, this study contributes to work on digital pragmatics and the sociolinguistics of social media by illustrating how small pragmatic items can circulate across languages and become routinised in platform genres (Schellewald 2021). It also reinforces a methodological point: frequency patterns are useful, but they only become meaningful when read alongside context. For this reason, I combined distributional observations with close, clip-by-clip interpretation of marker placement and function.

The study has obvious limitations. The corpus is small and platform-specific, and TikTok's trend cycles and visibility mechanisms may shape what is easiest to collect and observe. Future research could expand the corpus, diversify genres and creators, compare platforms, and add longitudinal sampling to see whether these routines stabilise or shift over time. Interviews would

also be valuable, because they could capture creators' own norms and motivations without relying on inference from public clips alone.

In conclusion, the thesis suggests that English discourse markers in Romanian TikTok speech reflect language-contact dynamics under digitally mediated conditions. In the clips analysed here, these markers act as economical tools for organising talk, marking stance, and building audience alignment, functions that are especially useful in short-form, fast-paced communication.

## Selected Bibliography

1. Avram, Mioara, 1997, “Anglicismele în limba română actuală”, in *Conferințele Academiei Române: Ciclul Limba română și relațiile ei cu istoria și cultura românilor*, Editura Academiei Române, București;
2. Bărbuleț, Gabriel, 2013, “Social Media – A Pragmatic Approach: Contexts & Implicatures”, in *Procedia - Social and Behavioral Sciences* 83, pp. 422-426;
3. Barker, George, C., 1945, “The Social Functions of Language”, in *ETC: A Review of General Semantics*, vol. 2, no. 4, pp. 228–34, *JSTOR*, <http://www.jstor.org/stable/42575938>, accessed on July 30<sup>th</sup>, 2024;
4. Blommaert, Jan, 2011, “Language and Superdiversity”, in *Diversities*, vol. 13, No. 2, pp. 1-21;
5. Brown, Gillian, Yule, George, 2020 [1983], *Discourse Analysis*, Cambridge University Press, United Kingdom;
6. Bucholtz, Mary, Hall, Kira, 2004, “Language and Identity”, in *A Companion to Linguistic Anthropology*, chapter 16, Blackwell, pp. 369-394;
7. Chung, Siaw-Fong, Liu, Hui-Wen, 2023, “The Pragmatics of Social Media”, in *International Review of Pragmatics*, 15, pp. 165-167;
8. Crystal, David, 2003, *English as a Second Language*, second edition, Cambridge University Press, New York;
9. Crystal, David, 2004, *Language and the Internet*, Cambridge University Press, United Kingdom;
10. Cojocaru, Valentina, 2020, “English discourse markers in spoken Romanian: pragmatic borrowings or a code-switching phenomenon?”, in *RRL*, LXV, nr. 3, Editura Academiei, București, pp. 261-271;
11. Cummins, Jim, 2017, “Teaching Minoritized Students: Are Additive Approaches Legitimate?”, in *Harvard Educational Review*, vol. 87, No. 3, Cambridge, pp. 404-425;
12. Fraser, Bruce, 1996, “Pragmatic Markers”, in *Pragmatics*, Vol. 6, No. 2, Jun., Boston University, pp. 322-343, [https://www.researchgate.net/publication/43647684\\_Pragmatic\\_Markers](https://www.researchgate.net/publication/43647684_Pragmatic_Markers), accessed on July 30<sup>th</sup>, 2024;

13. Grosjean, François, 2010, *Bilingual: Life and Reality*, Harvard University Press, Cambridge;
14. Gumperz, J. J., 1968, “The Speech Community”, in *D. L. Sills (Ed.), International Encyclopedia of the Social Sciences*, Macmillan, New York;
15. Halliday, M. A. K., 1978, *Language as Social Semiotic: the Social Interpretation of Language and Meaning*, University Park Press, Baltimore;
16. Hymes, D.H., 1974, *Foundations of Sociolinguistics*, University of Pennsylvania Press, Philadelphia;
17. Hudson, R. A., 1996, *Sociolinguistics*, Cambridge University Press, United Kingdom;
18. Jakobson, Roman, 1960, “Linguistics and Poetics”, in *Sebeok, T., ed., Style in Language*, MA: M.I.T. Press, Cambridge, pp. 350-377;
19. Leppänen, Sirpa, Peuronen, Saija, Westinen, Elina, 2018, “Superdiversity Perspective and the Sociolinguistics of Social Media”, in *The Routledge Handbook of Language and Superdiversity*, 1<sup>st</sup> edition, Routledge, London;
20. Niculescu-Gorpin, Anabella-Gloria, Vasileanu, Monica, 2020, “Romanian anglicisms: from fully-fledged lexical items to discourse markers”, in *Revue roumaine de linguistique*, LXV (3), Editura Academiei, București, pp. 273-284;
21. Oltean, Ștefan, 1977, “Considerații asupra naturii limbajului uzual”, in *Cercetări de lingvistică*, XXII, 1, Cluj-Napoca, Editura Academiei, pp. 43-47;
22. Poplack, Shana, 1980, “Sometimes I’ll start a sentence in Spanish Y TERMINO EN ESPAÑOL: toward a typology of code-switching”, in *Linguistics*, 18, 581–618;
23. Robertson, Roland, 1992, *Globalization: Social Theory and Global Culture*, Sage, London;
24. Roudometof, Victor, 2016, “Theorizing Glocalization: Three Interpretations”, in *European Journal of Social Theory*, vol. 19 (3), pp. 391-408;
25. Schellewald, Andreas, 2021, “Communicative Forms on TikTok: Perspectives from Digital Ethnography”, in *International Journal of Communication*, vol. 15, pp. 1437–1457;
26. Schiffrin, Deborah, 1987, *Discourse Markers*, Cambridge University Press, New York;
27. Vennemann, Theo, 2011, “English as a Contact Language: Typology and Comparison”, in *Anglia - Zeitschrift für englische Philologie*, December, England, pp. 217-257.