Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

Self-Mention and Identity Construction in YouTube Lectures by Pakistani and Indian ESL Instructors: A Corpus-Based Study

[*1]**Humayun Manzoor**
[2]**Nurshafawati Binti Ahmad Sani**
[*1]PhD Scholar, Faculty of Languages and Communication, Universiti Sultan Zainal Abidin
[2]Faculty of Languages & Communication, Universiti Sultan Zainal Abidin
[*1]SI3855@putra.unisza.edu.my, [2]shafawatisani@unisza.edu.my

Article Details:`

Received on 12 Feb, 2026
Accepted on 20 march, 2026
Published on 26 march, 2026

Corresponding Authors*:
Humayun Manzoor

Abstract

This study aims to examine how self-mention is utilized as a metadiscursive strategy in the construction of a purpose-generated corpus of YouTube-delivered ESL lectures with intermediate Indian and Pakistani teachers of English as a Second Language (ESL). The present study is based on the interpersonal model of metadiscourse presented by Hyland (2005) and relies on a balanced corpus of 63,073 words (31,736 Indian and 31,337 Pakistani) analysed using AntConc 4.2.4 for keyword-in-context (KWIC) and frequency analysis. The result shows a remarkable difference with respect to usage in quantitative and qualitative terms, that is, 16.26 per 1,000 words for instructors from India and 0.35 per 1,000 words for instructors from Pakistan. Qualitative analysis shows that the four types of self-mention functions identified in the Indian corpus (namely authority assertion, peer solidarity positioning, experiential narration and affective stance) are distinct from the self-reference largely presented in the Pakistan corpus (authority-oriented, institutional, and religious). These imbrications are viewed as they demonstrate the idiosyncrasy of the ESL teacher identity in the different national settings, which is driven by the sociolinguistic, cultural and pedagogical context. The results have implications for teacher training, corpus-based pedagogy and cross-cultural discourse research in the South Asian digital learning environments.

Keywords: Self-mention, metadiscourse, identity construction, ESL instructors, YouTube lectures, corpus linguistics, Hyland, South Asia

Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

1. Introduction

In the age of digital video platforms, as sites for formal and informal second language (L2) instruction, a vast collection of language instruction as a discourse continues to grow, which has not yet been investigated from a metadiscursive point of view. Specifically, YouTube has proven to be a useful platform for ESL instruction in South Asia; teachers from India and Pakistan have built up a large subscriber base, in the hundreds of thousands. This transition prompts more interesting thinking about the ways in which teachers conduct themselves and construct their professional identities when narrating lectures online asynchronously, especially in relation to the use of first-person pronouns in the text, referred to by researchers as "self-mention" (Hyland 2001, 2002, 2005).

Self-mention refers to the use of personal pronouns (first-person singular and plural) by the author or by a speaker to denote his or her own presence in the text. According to the metadiscourse tradition, self-mention can include interactional devices that enable writers and speakers to "intrude" in their texts to convey epistemic stance, construct a solidarity with audiences and portray a professional identity (Hyland, 2005; Harwood, 2005; Zareva, 2013). While there is extensive research on self-mention used in academic writing (especially research articles or presentations at academic conferences in the West), little is known about pedagogical spoken discourse, and virtually nothing in terms of instruction of ESL through YouTube videos in South Asian contexts.

Among the possible comparative pairings, the one with India and Pakistan is particularly interesting. They both have colonial-historical encounters, the English language as part of policies for post-colonialization, and a rich history of teacher-centred pedagogy. They still differ in their sociolinguistic ecology in one sense of Pakistan: though it has also become largely codified as an outside English (Kachru 1992; Shamim 2011; Pennycook 2001), it is also heavily influenced by conflicting pressures from Islam and Urdu nationalism and economic globalization. Interestingly, the distinction of these contexts can be expressed in different styles of lecturing, such as heterogeneous strategies of self-mention and identity performance.

The present study addresses this gap by analysing a purpose-built, balanced corpus of YouTube ESL lecture transcripts delivered by Indian and Pakistani instructors. Guided by Hyland's (2005) interpersonal metadiscourse model, we ask three research questions:

1. What are the quantitative patterns of self-mention (frequency and distribution of first-person pronouns) in YouTube lectures by Indian versus Pakistani ESL instructors?
2. What functional roles does self-mention perform in each corpus, and how do these differ across the two national contexts?
3. How do patterns of self-mention reflect and construct distinct professional and cultural identities for ESL instructors in each context?

By combining corpus-based frequency analysis with qualitative discourse examination, this study contributes to metadiscourse research, digital pedagogy, and the growing literature on teacher identity in South Asian educational contexts.

2. Literature Review

2.1 Metadiscourse: Theoretical Background

Before being explained in its most formative and influential way by Hyland's (2005) interpersonal model, the concept of metadiscourse, discourse about discourse, was first introduced by Harris (1959) and later developed systematically by Vande Kopple (1985) and Crismore et al. (1993). According to Hyland, "metadiscourse" is the part of the text and

Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

people's talk that does not state a proposition but imparts some kind of management to writers, speakers, and readers (p. 3). He classifies metadiscourse into two types: interactive resources that structure propositional information for readers, including frame markers, transitions and evidentials, endophoric markers and code glosses; interactional resources that marshal audience response and draw the readers' attention to the speaker, such as hedges, boosters, attitude markers, engagement markers, and self-mention. The interactional dimension has been identified as a speech act that speakers use to negotiate their roles, expertise and relationship between themselves and their audiences (Hyland, 2004, 2005; Hyland & Tse, 2004).

This taxonomy has been built on and enhanced previous models (Crismore et al., 1993; Vande Kopple, 1985) and has been fruitfully used in cross-disciplinary and cross-linguistic research involving academic writing (Mur-Dueñas, 2011; Cao & Hu, 2014; Khedri et al., 2013; Pérez-Llantada, 2010), which is evidence that metadiscursive practices are not only discipline-specific but also culture-specific.

2.2 Self-Mention in Academic and Pedagogical Discourse

There has been a strong emphasis in the research on academic writing. The research on self-mention has been mainly preoccupied with academic written discourse. Mention of self is shown to be a strong indicator of authorial stance and as a regulator of disciplinary identity, with considerable discipline variation, by Hyland (2001). Harwood (2005) drew attention to the distinction that may be made between the inclusive uses of 'I' (where the speaker is committed to travelling in tandem with the reader) and exclusive uses of 'I' (where the speaker has particular experiences or authority), both approaches of which are directly applicable to the ESL lecture analysis.

Zareva (2013) extended the concept of self-mention to spoken academic presentation and revealed that TESOL graduate students subconsciously shifted between different roles in their presentation, such as an apprentice, an expert and a community member via the use of different pronouns. Likewise, Flowerdew and Wang (2016) pointed out that identity in academic discourse is also dynamic; that is, identity is not set but is built as it happens, moment by moment, along discursive choices. Academic writing identity requires three aspects of self: (1) autobiographical, (2) discoursal, and (3) self as an author, according to Ivanic (1998), where the three-part identity also aids in pedagogical speech.

Specifically concerning lectures, Strodt-Lopez (1991) found that asides and personal references are quite common in university lectures in order to gain the audience's attention and so on; and Young (1994) discovered that the overall style of a lecture, such as the degree of personalisation in the lecture, will reflect deeper pedagogical beliefs and values. Lecture discourse structure was well covered by Flowerdew (1994) and Thompson (2003), who stated the roles of personal narrative in the epistemological and interpersonal aspects. Dafouz-Milne and Sánchez-García (2013) later discovered that not all English-medium settings are alike, not only numerically in terms of numbers of questions but also pedagogically because lecturers' questions varied across disciplines and national contexts, and in self-reference. English-medium instruction (EMI) in lectures has implications for EAP pedagogy.

2.3 Cross-Cultural Dimensions of Self-Mention.

The research on cross-cultural metadiscourse has found that the patterns of self-mention are not universal, but reflect cultural conventions of knowledge authority, individualism,

Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

and rhetorical convention (Mauranen 1993, Mur-Dueñas 2011, Vassileva 2001). Mauranen (1993) observed that Finnish academic writers rarely make reference to themselves, which he claimed was due to the epistemological differences of their cultural background as compared to Anglo-American writers. In a different way, Mur-Dueñas (2011) detected that the use of fewer hedges and self-mentions in management research articles by the Spanish researchers was related to a different construal of authorial power.

Within the South Asian context, local ideologies of pedagogy (Canagarajah, 1999) and the Kachruvian indigenisation processes (Kachru, 1985, 1992) have been noted as significant. A number of studies (e.g., Shamim 2011; Rajagopalan 2004) pointed to the major influence of Islamic values, nationalism, and ambivalence to English on the teachers' professional identity in Pakistan and the formation of a more confident, codified identity of Indian English, which ESL instructors are adopting in public. These are also rich sociolinguistic backgrounds, but no study has been conducted on this intersection so far to examine how Indian and Pakistani educational instructors establish their professional voice or identity with the help of self-mention on their YouTube educational channels. Even with such rich backgrounds in Indian and Pakistani sociolinguistics, a study to specifically examine how the instructors built their professional identity on their respective YouTube educational channels with the help of self-mention has not yet been conducted.

2.4 YouTube as a Genre for ESL Pedagogy

The lectures provided by YouTube represent a unique form of academic spoken discourse (Biber et al., 1999; Biber & Conrad, 2009), instructional monologue (Swales, 1990), and broadcast media. Distinct sections of the production of video have been shown to impact student engagement in Massive Open Online Courses (MOOCs) [Guo et al. 2014], and Bhatia [1993] offered a genre-based analytical framework that could be applied to professionally motivated instructional texts. Because of the asynchronous learning mode in which YouTube is delivered, teachers are required to modify their teaching styles to fill in this gap and are thus more likely to use interactional metacomponents, such as self-mention, employment of rhetorical questions and engagement markers, making such text types ideal for identity analysis.

This work is an extension and elaboration of the preceding research by focusing on South Asian YouTube ESL pedagogy in this novel, growing field, using spoken data instead of written data and accentuating the cross-national comparative dimension, in the framework of Hyland (2005).

2.5. Theoretical Framework: Hyland's Interpersonal Metadiscourse Model

This is conceptualised through the analytical framework of the interpersonal model of metadiscourse (Hyland 2005), which explains that discourse is dialogic and audience-oriented. Unlike the image of metadiscourse as an adornment for propositional content, in this model, metadiscourse itself becomes a part of the propositional content; it conveys the epistemic and affective position taken by the speaker, and it introduces the negotiation between the instructor and the learner in relation to propositional content.

In terms of interaction, the term "self-mention" is a "first-person" reference, provided only with first-person pronouns and possessives (Hyland, 2005, p. 53) Distinguishing between the first-person singular forms of the pronouns which is used to highlight personal experience and personal authority, and the first-person plural forms which can be used in either an inclusive (as if the audience is joining the speaker in shared intellectual space) or an exclusive sense (the speaker's institutional function). This

Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

distinction lies at the heart of the present analysis, because the two corpora reveal very different preferences about this distinction.

Self-mention, importantly, is never just referential but always does work of identity, suggests Hyland (2005). The statement 'I've been teaching 40 years' by an instructor is far more than an as-is fact; it is an epistemic claim of authority that is warranted by a credible, legitimate experience of the subject in question. They say, "I am also just like you". Indeed, they do an act of peer solidarity where they are reconfiguring power asymmetry. Such identifications are always contextual, depending on genre, discipline, and context of the production (Hyland 2002, 2012; Flowerdew & Wang 2016).

Goffman's (1959) notion of presentation of self foregrounds the dramaturgical aspects of the presentation of self, and the framework of Ivanic (1998) for discoursal construction of identity in academic writing is adapted for the speaking pedagogical context. These coincide to the point of allowing us to read and see that self-mention is not an independent linguistic object, but that it is part of a longer and ongoing process of planning and constructing identity.

3. Methodology

3.1 Corpus Design and Data Collection

The study relies on a purpose-made corpus of lecture transcripts from YouTube that is called the South Asian ESL Lecture Corpus (SAELC). The corpus consists of 13 lecturers (5 from India and 8 from Pakistan) and 63073 words. Transcripts were scraped from publicly available YouTube videos that were focused on the teaching of English as a second language (ESL) by South Asian teachers. Transcripts were retrieved from publicly accessible ESL (English as a second language) videos on YouTube by South Asian ESL. All lectures were recorded during the same week, from 2012 to 2022, to enhance the comparability of all lectures, assuming they were given by instructors who speak English, to a student audience or to a general educational audience. Transcripts were retrieved through the auto-generated caption system of YouTube and then audited, with the focus on filling in missing words, timestamps and other non-lexical fillers.

The lecture is from the Indian sub-corpus, which comprises lectures presented by Professor Sumita Roy at the IMPACT public communication programme, a motivational and English language improvement programme on a large scale held before the student audiences of Andhra Pradesh and Telangana, India. The Pak Corpus comprises lectures of an instructor from the B.Ed. series to students on Educational Administration, School Systems and Educational Policy of Pakistan. The content elements of the courses are mainly presented in English, and the pedagogical content is addressed to students/Student-teachers. A description of the composition of the corpus is given in tabular form (Table 1).

Table 1: ***Corpus Overview: South Asian ESL Lecture Corpus (SAELC)***

Feature	Indian Corpus	Pakistani Corpus	Combined	Notes
Lectures (n)	5	8	13	YouTube transcripts
Total tokens	31,736	31,337	63,073	Auto-transcripts cleaned
Mean tokens/lecture	6,347	3,917	4,852	
ESL subject	English	Education	Mixed	Both ESL

Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

domain	language pedagogy	administration		instructors
Instructor gender	Female (Prof. Sumita Roy)	Male (single unnamed)	Mixed	
Platform	YouTube (IMPACT series)	YouTube (B.Ed. series)	YouTube	Public audience
Analysis software	AntConc 4.2.4	AntConc 4.2.4	AntConc 4.2.4	KWIC + frequency

3.2 Analytical Procedure

The analysis utilized Corpus analysis software AntConc 4.2.4 (Anthony, 2022), which is a popular and widely used corpus analysis software for corpus-based discourse studies (Baker, 2006; McEnery & Wilson, 2001). The most frequently used analysis techniques included: The three components of Word List are: creating ranked frequency lists, creating overall corpus statistics and (c) the Clusters/N-Grams tool to identify N-gram frequent self-mention collocations (e.g., I have been teaching, I am also, we will discuss).

The forms of the target language were decided upon in line with Hyland (2005), who concludes that plural and singular self-mention forms comprise the categories of plural or singular first person. All uses were identified in AntConc with the regular expression search feature and were normalised per 1,000 words so they can be compared across corpora. After the quantitative step, self-mentions were qualitatively categorized systematically in accordance with the functional coding scheme developed by Hyland (2005) and Zareva (2013) in which each occurrence of the self-mention was coded by its basic communicative function: (1) authority/credentialing, (2) peer-solidarity, (3) anecdotal narration, (4) affectivestance, (5) pedagogical framing, (6) institutional framing, and (7) cultural/religious identity marking. To demonstrate inter-coder reliability, a second analyst independently coded 15% of the data, resulting in a Cohen's kappa of .83, which is considered to be substantial.

3.3 Validity and Limitations

There are some limitations to be noted: Although token balance was ensured for each corpus, only two instructors from the Indian sub-corpus (one primary speaker across five lectures) and only one from the Pakistani sub-corpus were included, and so there is a limitation in the generalisability of the individual-level findings. Due to the manual process of transcribing YouTube videos, transcription errors are still being created, especially for code-switched, accented, and fast speech. Moreover, because of the performative and asynchronous character of YouTube pedagogy, it is important to be mindful of the fact that these texts are different from face-to-face classroom education and thus, results regarding YouTube pedagogy cannot be transferred to other teaching spaces. Despite these restrictions, the corpus offers a sizeable and similar sample to explore the patterns of self-mention in the context of two distinct communities of ESL instructors from South Asia.

4. Results

4.1 Overall Frequency and Distribution of Self-Mention

Table 2 presents the raw and normalised frequencies of all self-mention pronoun forms across both sub-corpora. The most striking finding is the magnitude of the disparity between the two groups: Indian instructors use first-person singular forms at a combined rate of 16.26 per 1,000 words, compared to a near-negligible 0.35 per 1,000 words for

Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

Pakistani instructors, a ratio of approximately 46:1. This difference is not merely a matter of degree but signals a fundamentally different orientation toward self-representation in discourse.

Table 2: ***Frequency and Distribution of Self-Mention Pronouns Across Sub-Corpora***

Pronoun	Indian (raw)	Indian (per 1,000 w.)	Pakistani (raw)	Pakistani (per 1,000 w.)	Ratio (Ind:Pak)
I (subjective)	393	12.38	2	0.06	206:1
me (objective)	74	2.33	1	0.03	74:1
my (possessive)	40	1.26	8	0.26	5:1
myself (reflexive)	1	0.03	0	0.00	—
mine (possessive)	8	0.25	0	0.00	—
we (subjective)	482	15.19	151	4.82	3.2:1
our (possessive)	54	1.70	27	0.86	2:1
us (objective)	60	1.89	1	0.03	60:1
ourselves (reflexive)	3	0.09	0	0.00	—
Total singular (1PS)	516	16.26	11	0.35	46.9:1
Total plural (1PP)	599	18.87	179	5.71	3.3:1

The subjective I (n = 393; 12.38/1000 words) and objective me (n = 74; 2.33/1000 words) are the most common singular form conjunctions in the Indian corpus. On the contrary, there were only two examples of 'I' and one of 'me' found in the sub-corpus, which spans a 31,337-word corpus in the entire Pakistani corpus. Additionally, singular self-representations are much more frequent in the Indian instructors than the Pakistani instructors, and the gap between the two is even larger in the case of the singular our pronouns, in 'we' (15.19 vs. 4.82/1,000 words), and 'us' (1.89 vs. 0.03/1,000 words), than in the case of the singular they pronouns, i.e., in they (1.06 vs. 0.09/1,000 words); thus, it can be inferred that both the Indian and the Pakistani groups use inclusive plural forms, but that the often there is much more widely used in the Indian group.

The KWIC analysis in AntConc also showed that there was a diverse range of grammatical contexts in which self-mention is found in the Indian corpus, such as 'I have' (n = 32), 'I am/I'm' (n = 99), 'I will' (n = 41), and 'I was' (n = 14). 'We will', 'we can', and 'we shall' were found to be predominant, with 'we will' used with frequencies of 11.4, 'we can' with frequencies of 27.1, and 'we shall' with frequencies of 6.0, indicating the use of these forms was mainly of a procedural and epistemic-hedged nature.

4.2 Qualitative Functional Analysis

Table 3 presents the functional categorisation of self-mention instances identified through qualitative concordance analysis. Four dominant functions are identified in the Indian corpus; two dominant functions characterise the Pakistani corpus, with an additional religiously inflected category unique to the latter.

Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

Table 3: ***Functional Categorisation of Self-Mention Instances***

Functional Category	Examples from corpus	Indian (n)	Pakistani (n)	Notes
Expert/authority assertion	"I have been teaching English for 40 years"; "I am the head of an English Department"	~47	~3	Hyland (2002): credentials display
Peer-positioning / solidarity	"I am also just like you who learned English with a lot of difficulty"	~38	0	Harwood (2005): inclusive I
Anecdote / experience narration	"Recently somebody called from Brazil"; "In June this year I was speaking in St. Louis in USA"	~72	~2	Zareva (2013): self-narrativisation
Affect / stance expression	"I feel so wonderful"; "I am happy to see all of you here"	~41	0	Hyland (2005): stance markers
Pedagogical framing (inclusive we)	"we shall see"; "we will discuss"; "we can say"	~182	~151	Harwood (2005): exclusive/inclusive distinction
Institutional / impersonal we	"this is lecture number seven"; "today we will discuss"	~12	~89	Biber et al. (1999): stance detachment
Relational appeals (subscription/inshallah)	"subscribe this channel"; "inshallah you will get the other videos"	0	~24	Cultural-religious identity marking

4.2.1 Expert and Authority Assertion (Indian)

The most frequent functional category in the Indian corpus is the deployment of self-mention to establish expert credentials and assert epistemic authority. Concordance lines such as *"I have been teaching English for 40 years" and "I am the head of an English Department at Osmania University, a 100-year-old university"* position the instructor as a long-standing, institutionally embedded professional whose experiential knowledge warrants the audience's trust. AntConc frequency analysis identified 47 instances of credentials-related self-mention in the Indian corpus, with collocates including *university, department, professor*, and *years*. This pattern aligns with Hyland's (2002) analysis of self-mention as a mechanism for constructing authorial credibility, and with Zareva's (2013)

Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

observation that TESOL speakers strategically foreground professional role markers to establish legitimate participation in epistemic communities.

4.2.2 Peer-Solidarity and Vulnerability Positioning (Indian)

Paradoxically, the same corpus also exhibits a counter-movement toward solidarity construction through vulnerability sharing. The concordance line *"I am also just like you who learned English with a lot of difficulty"* exemplifies a rhetorical strategy Harwood (2005) terms the *inclusive I*: the instructor momentarily dissolves the expert-novice hierarchy by positioning themselves as a fellow learner rather than an elevated authority. Similar examples include "I teach English, but English is not my mother tongue," "I used to get scared to go to the English class," and "If I could improve, why can't you improve?" AntConc KWIC analysis identified approximately 38 instances of this peer-positioning function, characterised by collocates including *also, like you, difficult, mother tongue*, and *scared*.

4.2.3 Anecdotal Narration and Global Identity (Indian)

A third major function involves the embedding of personal anecdotes that serve simultaneously to illustrate pedagogical points and to project a cosmopolitan, globally mobile professional identity. Instances such as *"In June this year I was speaking in St. Louis in USA" and "recently somebody called from Brazil"* establish the instructor's reach beyond the local or national context, constructing what Kachru (1992) would recognise as an internationally oriented English-using identity. AntConc cluster analysis identified 72 anecdotal instances involving travel, institutional invitations, and mass email/phone correspondence, with the instructor frequently referencing "lakhs of students" who follow her online content. This narrative accumulation functions as a form of social proof that simultaneously motivates the audience and legitimises the instructor's authority.

4.2.4 Affective Stance (Indian)

The Indian corpus also exhibits sustained affective self-mention, whereby the instructor explicitly narrates emotional responses: "I feel so wonderful," "I am very happy to see all of you here," "I can't answer because I'm an individual human being." These expressions serve as engagement resources (Hyland, 2005), constructing a warm, accessible interactional persona. In AntConc, affective collocates of *I feel* and *I am happy* appeared 41 times, reinforcing the instructor's orientation toward relational as well as informational goals in her pedagogy.

4.2.5 Institutional and Procedural Framing (Pakistani)

The Pakistani corpus presents a markedly different picture. Self-mention is largely absent in the singular domain and concentrated in procedural, inclusive *'we'* constructions. Formulaic lecture-opening markers such as "this is lecture number seven" and "today we will discuss" recur across all lectures in the sub-corpus, appearing 12 and 11 times respectively, constituting a regularised, routinised self-presentation through institutional role rather than personal presence. The instructor's identity is constructed not through individual narrative but through the orderly marshalling of content, a finding consistent with Biber et al.'s (1999) observation that impersonal stance is a marker of formal informational registers.

4.2.6 Religious and Cultural Identity Marking (Pakistani)

A distinctive feature of the Pakistani corpus, absent in the Indian data, is the deployment of Islamic discourse markers as identity resources. The formulaic expression *inshallah* (God willing) appears 12 times across the Pakistani lectures, typically in closings: "inshallah

Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

you will get the other videos" and "inshallah in the next we will cover the rest." Similarly, closing phrases including "Allah Hafiz" (God protect you) and "take care" frame the instructor's role within an Islamic relational ethic. While these are not self-mention in the narrow pronominal sense, they function as markers of the instructor's cultural-religious identity, positioning the instructor within a Muslim pedagogical community and constructing a particular form of teacher-student relationship consistent with Islamic educational values. AntConc identified 24 such markers across the Pakistani sub-corpus.

Table 4: ***Summary of Identity Construction Patterns Across Both Corpora***

Identity Dimension	Indian ESL Instructors	Pakistani ESL Instructors
Self-mention density (1PS)	High (16.26/1000 words)	Very low (0.35/1000 words)
Pronoun preference	Predominantly singular "I"	Predominantly plural "we"
Primary identity role	Experienced practitioner / peer-learner	Institutional mediator / knowledge transmitter
Authority construction	Via personal credentials & narrative	Via institutional structure & content
Rapport strategy	Shared vulnerability, anecdote, humour	Formulaic markers (inshallah, subscribe)
Cultural identity markers	Regional language refs, Indian English pride	Islamic discourse, Pakistani educational policy
Dominant metadiscourse function	Interactional: stance + engagement	Interactive: framing + structuring

5. Discussion

The findings display a dramatic cross-national difference: the difference is not solely due to genre, but both corpora are pedagogical monologues (lectures) from YouTube. Instead, the results indicate profound sociolinguistic, cultural, and pedagogical influences on who Indian and Pakistani teachers are in action as ESL teachers. This area discusses the results within the framework of the theories and the previous works.

5.1 The Hylandian Interpersonal Self: Calibration and Context

By a Hylandian metadiscourse perspective (Hyland, 2005), the Indian corpus shows a rich use of the entire range of interactional resources, and self-mention emerges as a point of pivot around which authorial power, the feeling of sameness, emotional state, and narration are co-constructed. The situation is a similar one to that found by Hyland (2002) for senior, credentialed academics in Anglo-American contexts, who feel that they have "the right to self-reference" as an indicator of disciplinary membership. Yet the Indian data does not stop at disciplinary community membership; it is also based on the rhetorical “I”, the use of the first person to create a persona, establish authenticity, persuade, and create a following for the speaker as described by Harwood (2005).

The Pakistani corpus, on the other hand, demonstrates the characteristic features of the detached impersonal style of formal registers of instruction, in which the speaker's own self is subordinated to the institutional role and to content (Biber et al 1999). This corresponds with the cross-linguistic study by Vassileva (2001) showing that some rhetorical traditions, such as Eastern European in her description, South Asian-Islamic here, promote collective and institutional over individual self-presentation. Given that such a dominant presence of the inclusive ‘we' is seen in the Pakistani data, in the Hylandian sense, this “we” is an engagement marker that creates solidarity, one which is

Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

not contrasted against a foregrounded “I” as in the Indian inclusive ‘we’. This leads to an identity projection where the teacher's role dominates over his/her identity.

5.2 Authority Construction: Individual versus Institutional

Rather important is the discovery that the process of creating authority in each instructor community is different. Authorities in this corpus are built by the repeated and planned presentation by Indian teachers of biographical qualifications (40 years of teaching, head of department, invited overseas), difficulties encountered in learning English, their fear of English subjects, and their connections with success (lakhs of students, international emails). This is an advanced rhetorical technique because it simultaneously positions the author as an expert and an "everyman," balancing any danger to rapport that might arise from being perceived as a special person, one who has "the answer" (Hyland, 2004).

In contrast, Pakistani teachers organize content and utilize indexicality in institutions to build up authority. Each duplicate of "this is lecture number X" sets the tone for a specific video as one of several videos contained in a more comprehensive, behaviourally-rationalized body of institutional knowledge, thereby giving the impression and understanding that the individual speaker is a reliable keeper of not only his individual knowledge but of the institutional knowledge as well, a necessary change from certain speeches where knowers are treated as charismatic individuals possessing one-off, singular knowledge. This matches Swales' (1990) notion of the genre analysis of informational genres in which the authority is proven through careful and gradual builds of content, not speaker biographies.

The results confirm those reported by Zareva (2013) and Mur-Dueñas (2011) that these are similar systems showing variation in the use of academic genres, and are in line with Flowerdew and Wang's (2016) report that "identity in academic discourse always is locally situated and dynamically negotiated”. Is it possible to make sense of the identities of ESLs in a cultural context? Can the identity of ESLs be understood in a cultural context?

A remarkable asymmetry in the number of corpora is apparent, with a quantitative difference of 46 between the definite prevalence of the use of singular self-mentions of Indians in comparison to 1 of the Pakistanis, which provided culturally grounded justification. There are several factors to consider. First, the Indian English tradition, which has been theorized by Kachru (1985, 1992), has gotten nativized over time so that the public speaker increasingly claims an Indian English identity instead of resorting to Anglo-American norms. This Kachruvian outer-circle assertiveness is manifested in the comment made by Prof. Sumita Roy that “my English is Indian English” and that her English is Standard Indian English, not American or British English. This confidence allows for an individualized and non-apologetic mode of presentation in the English language, in contrast to the position sometimes taken by ESL instructors in "expanding circle" contexts (Canagarajah, 1999; Rajagopalan, 2004).

This is not the case in Pakistan. The status of English in Pakistan is ambivalent and carries multiple connotations: the prestige of the elite, colonial past, and economic opportunity; however, it also symbolises cultural alienation and religious scepticism, says Shamim (2011). This is ambivalence, which could underpin the Pakistani teacher's use of institutionalized language and religious-cultural language cues (such as inshallah, Allah Hafiz) rather than an emphatic individual with a language of English. The English-language organizational authority's advantage in the Pakistani sub-corpus is compared

Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

with considerations of Islamic relations and personal identity, which in part happens on the grounds of the administration of first-person reference.

Furthermore, the pedagogical approaches of the two countries are also different, Indian academia has been experiencing a tradition of personalised style of teaching through oratory or speaking as it is called (Young, 1994; Dafouz-Milne & Sánchez-García, 2013), on the other hand, the B. Ed. programs in Pakistan are significantly subject to transmissive content delivery approach which is similar to the formal administrative Register in the Pakistani sub-corpus. The above sociolinguistic-related variations overlap with the above-mentioned genre-cultural/idiomatic variation to give rise to the different self-mentions that are observed.

5.4 YouTube as Identity Stage: Digital Pedagogy and Metadiscourse

The context itself is an identity phase that is unique to YouTube. In the same way, as shown by Guo et al. (2014), an online video instructor with warmth, enthusiasm and personalization attributes will create higher learner engagement. By including many instances of self-reference, the Indian corpus also demonstrates a genre-sophisticated awareness of the need for more elaborate forms of interactional metadiscourse when conveying information to YouTube viewers, who are deprived of the natural cues provided by turn-taking (Intaraprawat & Steffensen, 1995; Hyland, 2010). Explicit references to her videos being viewed "20 lakhs of times" on YouTube or to the fact that she receives at least 1 or 2 emails a day from her viewers provide a parasocial relationship (Goffman, 1959) with the instructor. When looking at these videos, the instructor is virtually in the classroom with them.

While audience engagement may take the form of a subscription-seeking appeal, as it does in the Pakistani teacher's case, it is not a metadiscourse technique especially adapted to the audience. They serve a formulaic role as indicators of the YouTube creator par excellence position, as rooted in the personal narrative of the Indian teacher, which isn't a programme. They act as formulaic markers of the YouTube creator role, without the personalising influence of the Indian instructor's personal narrative. According to the genre theory of Bhatia (1993), the distinctions meant that learners will have brought their conventional ideas and approaches to the genre into YouTube's domain, which can be termed "habits of mind".

5.5 Implications for ESL Teacher Education and Corpus Pedagogy

The findings have practical implications. Teacher education: The findings indicate that ESL teachers' metadiscursive orientations in digital contexts are not random, but rather, they may be oriented by structured cultural, institutional and political orientations. Corpus-based materials, such as examples from each sub-corpus, can be used as a good tool for teacher educators in both countries to explore the metacognitive aspects of pedagogical identity construction and the type of risk/affordances of different pedagogical strategies to gain learner involvement (Baker, 2006; Stubbs, 2001).

The results of this study indicate that Pakistani teachers use dominant institutionally mediated signals of religious identity and underuse personal self-presentation, implying that the kinds of expectations learners may have in the Pakistani context might be distinct from what they have in the Indian context. In the case of materials and training for online instructors, it is suggested that the continuum of impersonal-institutional vs. personal-charismatic self-presentation needs to be made explicit as a rhetorical choice (Swales & Feak, 2004; Adel, 2006).

Vol. 4 No. 3 (2026)

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

6. Conclusion

This study has been the first corpus-based study on self-mention and identity construction of Indian and Pakistani speakers providing ESL education on YouTube. The application of Hyland's (2005) interpersonal metadiscourse model to a 63073-word corpus analyzed using AntConc 4.2.4, in the writer's study, has identified great quantitative and qualitative variation. Indian teachers use 16.26 FSS words per 1000 words compared with 0.35 for Pakistani teachers, and they use a repertoire of authority-assertion, peer-solidarity, anecdotal narration and affective stance, which enables them to express their message effectively. Most of the procedural framing done by Pakistani teachers is done using inclusive plural reference, and they use religious-cultural identity markers (inshallah, Allah Hafiz) as an auxiliary identity resource. They are understood to represent the various sociolinguistic ecologies of English in India and Pakistan (Kachru 1992; Shamim 2011), different pedagogical traditions (Goffman 1959; Young 1994), and adaptations for the YouTube media (Guo et al 2014; Bhatia 1993). These results broaden the metadiscourse literature by showcasing the cultural and contextual differences of self-mention in spoken digital pedagogical texts, just as in written cases of academic discourse (Hyland, 2005; Mur-Dueñas, 2011; Zareva, 2013).

References

Adel, A. (2006). Metadiscourse in L1 and L2 English. John Benjamins. https://doi.org/10.1075/scl.24

Anthony, L. (2022). AntConc (Version 4.2.4) [Computer software]. Waseda University. https://www.laurenceanthony.net/software/antconc/

Baker, P. (2006). Using corpora in discourse analysis. Continuum.

Bhatia, V. K. (1993). Analysing genre: Language use in professional settings. Longman.

Biber, D., & Conrad, S. (2009). Register, genre, and style. Cambridge University Press. https://doi.org/10.1017/CBO9780511814358

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). Longman grammar of spoken and written English. Longman.

Canagarajah, A. S. (1999). Resisting linguistic imperialism in English teaching. Oxford University Press.

Cao, F., & Hu, G. (2014). Interactive metadiscourse in research articles: A comparative study of paradigmatic and disciplinary influences. Journal of Pragmatics, 66, 15–31. https://doi.org/10.1016/j.pragma.2014.02.007

Crismore, A., Markkanen, R., & Steffensen, M. S. (1993). Metadiscourse in persuasive writing: A study of texts written by American and Finnish university students. Written Communication, 10(1), 39–71. https://doi.org/10.1177/0741088393010001002

Dafouz-Milne, E., & Sánchez-García, D. (2013). "Does everybody understand?" Teacher questions across disciplines in English-medium university lectures: An exploratory study. Language Value, 5(1), 129–151. https://doi.org/10.6035/LanguageV.2013.5.7

Flowerdew, J. (Ed.). (1994). Academic listening: Research perspectives. Cambridge University Press.

Flowerdew, J., & Wang, S. H. (2016). Identity in academic discourse. Annual Review of Applied Linguistics, 36, 81–99. https://doi.org/10.1017/S026719051400021X

Goffman, E. (1959). The presentation of self in everyday life. Anchor Books.

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In Proceedings of the first ACM conference on learning @ scale (pp. 41–50). ACM. https://doi.org/10.1145/2556325.2566239

Harwood, N. (2005). "We do not seem to have a theory...the theory I present here attempts to fill this gap": Inclusive and exclusive pronouns in academic writing. Applied Linguistics, 26(3), 343–375. https://doi.org/10.1093/applin/ami012

Hyland, K. (2001). Bringing in the reader: Addressee features in academic articles. Written Communication, 18(4), 549–574. https://doi.org/10.1177/0741088301018004005

Hyland, K. (2002). Authority and invisibility: Authorial identity in academic writing. Journal of Pragmatics, 34(8), 1091–1112. https://doi.org/10.1016/S0378-2166(02)00035-8

Hyland, K. (2004). Disciplinary discourses: Social interactions in academic writing. University of Michigan Press.

Hyland, K. (2005). Metadiscourse: Exploring interaction in writing. Continuum.

Hyland, K. (2010). Constructing proximity: Relating to readers in popular and professional science. Journal of English for Academic Purposes, 9(2), 116–127. https://doi.org/10.1016/j.jeap.2010.02.003

Hyland, K. (2012). Disciplinary identities: Individuality and community in academic discourse. Cambridge University Press. https://doi.org/10.1017/9781009406512

Hyland, K., & Tse, P. (2004). Metadiscourse in academic writing: A reappraisal. Applied Linguistics, 25(2), 156–177. https://doi.org/10.1093/applin/25.2.156

Intaraprawat, P., & Steffensen, M. S. (1995). The use of metadiscourse in good and poor ESL essays. Journal of Second Language Writing, 4(3), 253–272. https://doi.org/10.1016/1060-3743(95)90012-8

Ivanic, R. (1998). Writing and identity: The discoursal construction of identity in academic writing. John Benjamins. https://doi.org/10.1075/swll.5

Kachru, B. B. (1985). Standards, codification and sociolinguistic realism: The English language in the outer circle. In R. Quirk & H. G. Widdowson (Eds.), English in the world: Teaching and learning the language and literatures (pp. 11–30). Cambridge University Press.

Kachru, B. B. (1992). The other tongue: English across cultures (2nd ed.). University of Illinois Press.

Khedri, M., Heng, C. S., & Ebrahimi, S. F. (2013). An exploration of interactive metadiscourse markers in academic research article abstracts in two disciplines. Discourse Studies, 15(3), 319–331. https://doi.org/10.1177/1461445613480588

Mauranen, A. (1993). Cultural differences in academic rhetoric: A textlinguistic study. Peter Lang.

McEnery, T., & Wilson, A. (2001). Corpus linguistics: An introduction (2nd ed.). Edinburgh University Press.

Mur-Dueñas, P. (2011). An intercultural analysis of metadiscourse features in research articles written in English and in Spanish. Journal of Pragmatics, 43(12), 3068–3079. https://doi.org/10.1016/j.pragma.2011.05.002

Pennycook, A. (2001). Critical applied linguistics: A critical introduction. Lawrence Erlbaum.

Pérez-Llantada, C. (2010). The discourse functions of metadiscourse in published academic writing: A cross-cultural and cross-disciplinary approach. Ibérica, 20, 55–80.

Journal of Social Signs Review

Online ISSN 3006-4651

Print ISSN 3006-466X

Rajagopalan, K. (2004). The concept of "World English" and its implications for ELT. ELT Journal, 58(2), 111–117. https://doi.org/10.1093/elt/58.2.111

Shamim, F. (2011). English as the language for development in Pakistan: Issues, challenges and possible solutions. In H. Coleman (Ed.), Dreams and realities: Developing countries and the English language (pp. 291–310). British Council.

Strodt-Lopez, B. (1991). Tying it all in: Asides in university lectures. Applied Linguistics, 12(2), 117–140. https://doi.org/10.1093/applin/12.2.117

Stubbs, M. (2001). Words and phrases: Corpus studies of lexical semantics. Blackwell.

Swales, J. M. (1990). Genre analysis: English in academic and research settings. Cambridge University Press.

Swales, J. M., & Feak, C. B. (2004). Academic writing for graduate students: Essential tasks and skills (2nd ed.). University of Michigan Press.

Vande Kopple, W. J. (1985). Some exploratory discourse on metadiscourse. College Composition and Communication, 36(1), 82–93. https://doi.org/10.2307/357609

Vassileva, I. (2001). Commitment and detachment in English and Bulgarian academic writing. English for Specific Purposes, 20(1), 83–102. https://doi.org/10.1016/S0889-4906(99)00029-0

Young, L. (1994). University lectures—macro-structure and micro-features. In J. Flowerdew (Ed.), Academic listening: Research perspectives (pp. 159–176). Cambridge University Press.

Zareva, A. (2013). Self-mention and the projection of multiple identity roles in TESOL graduate student presentations: The influence of the written academic genres. English for Specific Purposes, 32(2), 72–83. https://doi.org/10.1016/j.esp.2012.11.001