What Do LIWC Features Capture in Multilingual Classroom Talk? A Systemic Functional Linguistic Analysis

Kaycie Barron
University of Wisconsin–Madison
kdbarron@wisc.edu

Ananya Ganesh
University of Wisconsin–Madison
aganesh27@wisc.edu

Shamya Karumbaiah
University of Wisconsin–Madison
shamya.karumbaiah@wisc.edu

Mariana Castro
University of Wisconsin–Madison
mcastro@wisc.edu

ABSTRACT

Classroom discourse among multilingual learners (MLLs) routinely integrates cognitive, interpersonal, and organizational work, yet scalable discourse analytics in educational data mining (EDM) rely on linguistic features whose relationship to these meanings is often unclear. This study examines how psycholinguistic features derived from Linguistic Inquiry and Word Count (LIWC) correspond to expert interpretations of meaning in multilingual small-group science discourse. Grounded in Systemic Functional Linguistics (SFL), we analyze associations between LIWC categories and human annotations of the ideational and interpersonal metafunctions across ten classroom transcripts. Using three complementary analyses, we evaluate empirical alignment patterns and theory-driven predictions. Results show that LIWC features align most consistently with interpersonal meaning, capturing stance and social positioning, whereas ideational meaning is reflected more selectively in measures of elaboration and in references to participants and actions. These findings clarify what dimensions of classroom meaning are accessible as LIWC-based representations and highlight important limitations of psycholinguistic features for modeling learning-relevant discourse in multilingual settings.

Keywords

Educational Data Mining, Collaborative Learning, Multilingual Discourse, Systemic Functional Linguistics, Interpretability

1. INTRODUCTION

Classroom discourse plays a central role in collaborative learning by shaping how multilingual learners (MLLs) construct ideas, negotiate understanding, and coordinate participation during instruction. In multilingual classrooms, these processes are often realized through translanguaging practices in which learners flexibly draw on multiple linguistic resources to foster joint sense-making and participation [2, 16]. Prior research shows that student talk in small-group STEM activities routinely serves multiple communicative functions simultaneously, such as combining problem solving, interpersonal positioning, and discourse organization within a single spoken utterance [20, 5, 16].

Educational data mining (EDM) has increasingly turned to discourse analysis to study classroom interaction and learning processes at scale [25, 6]. This approach offers powerful tools for characterizing multilingual meaning-making and collaborative processes across large datasets. However, many commonly used linguistic representations rely on surface-level features whose relationship to learning-relevant discourse functions remains underspecified, particularly in multilingual settings where it does not conform to monolingual or standardized norms [3, 12]. For instance, a student might say, "No sé... [I don't know] but I think it's because it evaporates faster," where the phrase may register pronoun use, negation, or cognitive markers in surface feature models, while simultaneously contributing to collaborative hypothesis building that may not be fully represented.

Linguistic Inquiry and Word Count (LIWC) provides one widely adopted and interpretable approach to discourse analysis by quantifying psycholinguistic dimensions such as cognitive processing, affective tone, and social orientation [27, 22]. LIWC features have been used extensively in educational research to examine student engagement, performance, and experience, including in STEM learning contexts [23, 17]. Despite its widespread use, LIWC is often applied without an explicit theory of discourse and assumes monolingualism, leaving open questions about what kinds of communicative work these features actually reflect in linguistically diverse classroom interactions. In addition, MLLs routinely draw on multiple languages for both conceptual sense-making and social interaction, even when English is the official medium of instruction [8, 21]. Approaches that rely on language-specific dictionaries risk under-representing how meaning is distributed across languages, with implications for how discourse is interpreted and used in downstream models of learning.

Kaycie Barron, Ananya Ganesh, Shamya Karumbaiah, and Mariana Castro. What Do LIWC Features Capture in Multilingual Classroom Talk? A Systemic Functional Linguistic Analysis. In Anthony Botelho, Maria Mercedes T. Rodrigo, Adish Singla, Hiroaki Ogata, Hyojeong So, and Young Hoan Cho (eds.) Proceedings of the 19th International Conference on Educational Data Mining, Seoul, Republic of Korea, June, 2026, pp. 388–399. International Educational Data Mining Society (2026).

https://doi.org/10.5281/zenodo.21040154

This study addresses this gap by examining how LIWC features align with human interpretations of discourse in K-12 multilingual science classrooms. Drawing on Systemic Functional Linguistics (SFL), which theorizes how meaning is realized through language across the ideational, interpersonal, and textual metafunctions [10, 19], we analyze the correspondence between select LIWC variables and expert-annotated metafunctions in small-group science discussions. Rather than treating linguistic features as predictors of downstream outcomes, we adopt a theory-informed analytic approach to estimate how LIWC's psycholinguistic variables represent different dimensions of meaning-making.

1.1 Multilingual Classroom Discourse

Research on classroom discourse has long emphasized that knowledge construction is achieved through interactional processes (i.e., explanation, questioning, critique, and uptake) instead of isolated individual contributions [20, 5]. In multilingual classrooms, these interactions are frequently characterized by translanguaging practices in which students coordinate meaning across and beyond languages and registers [7, 9]. Studies of multilingual STEM classrooms show that student contributions often integrate cognitive work, social alignment, and discourse organization within a single turn, reflecting the inherently multifunctional nature of collaborative sense-making [16].

While discourse analysis approaches have documented these patterns in detail, they typically rely on labor-intensive qualitative methods and small datasets, limiting their scalability for EDM applications [16]. This tension motivates the need for computational representations that remain sensitive to the functional complexity of classroom talk.

1.2 LIWC

LIWC is one of the most widely used dictionary-based tools for analyzing psychological, cognitive, and social dimensions of language. It operationalizes linguistic features as proportions of words associated with categories such as cognitive processes, affect, social relationships, and function words, as well as composite summary measures including Analytical Thinking, Clout, Authenticity, and Emotional Tone [27, 22].

In educational research, LIWC has been used to study how linguistic patterns relate to student outcomes and experience. For example, Lin, Yu, and Dowell [17] showed that cognitive and social LIWC features were differentially associated with performance and sense of belonging in online STEM courses. Other work by [23] showed through LIWC analysis that students' use of function words and analytic language patterns in their writing predicts academic performance.

Foundational LIWC research cautions that psycholinguistic categories do not map transparently onto communicative functions. The interpretation of any given LIWC feature depends on context, genre, and interactional setting, requiring theoretical grounding [27]. In multilingual classroom discourse, this limitation is compounded by LIWC's reliance on language-specific dictionaries. As MLLs fluidly integrate linguistic resources, analysis must either restrict data to one language or transform translanguaged discourse into monolingual representations. Both approaches risk distorting how meaning is distributed across languages and interaction, particularly in collaborative settings where ideational and interpersonal work are tightly integrated.

1.3 Systemic Functional Linguistics

SFL provides a principled framework for interpreting discourse by distinguishing among ideational and interpersonal metafunctions, which capture how speakers construe experience and negotiate social relations during interaction [19]. These dimensions have been widely used to analyze classroom talk and multilingual academic language practices, particularly in studies of collaborative learning and sense-making [26, 28].

- **Ideational**: construing experience, representing processes or phenomena, and elaborating content.
- **Interpersonal**: expressing stance, evaluation, uncertainty, and social alignment with others.

SFL also theorizes a *textual* metafunction concerned with cohesion, information flow, and discourse organization across extended stretches of discourse. However, textual meaning is primarily realized across turns and interactional sequences instead of isolated speaker turns. Because LIWC operates at the level of short utterances and relies on surface lexical and punctuation cues, it is poorly suited for capturing textual resources that develop thematically in dialogic classroom interaction. Accordingly, while textual meaning was annotated to characterize the multifunctional nature of classroom discourse, subsequent LIWC analyses in this study focus exclusively on ideational and interpersonal metafunctions.

2. METHODS

2.1 Dataset

The dataset consists of 204 utterances drawn from ten small-group science classroom transcripts collected across three Midwestern K–12 school districts. Transcripts were selected from longer classroom sessions as part of a broader study of multilingual interaction in science classrooms in Fall 2024 and Spring 2025. For the present analysis, we focus on pre-selected segments capturing focal episodes of collaborative sense-making and translanguaging.

Across these segments, 123 students participated. Classrooms varied in science content and linguistic demographics (four 6th-grade science classrooms, one 9th-grade biology classroom, and one 10th-grade chemistry classroom), and included MLLs with diverse language backgrounds (e.g., Spanish, Arabic) as well as monolingual English-speaking peers. The selected segments represent authentic small-group problem-solving interactions in which students collaboratively interpreted evidence, negotiated understanding, and coordinated action using their multilingual resources.

The data used in this study were collected as part of a broader research project approved by the Institutional Review Board (IRB) (ID: 2024-1029-CP004) at the University of Wisconsin-Madison. Informed consent was obtained from parents/guardians and assent from students in English and Spanish. All data were de-identified before analysis.

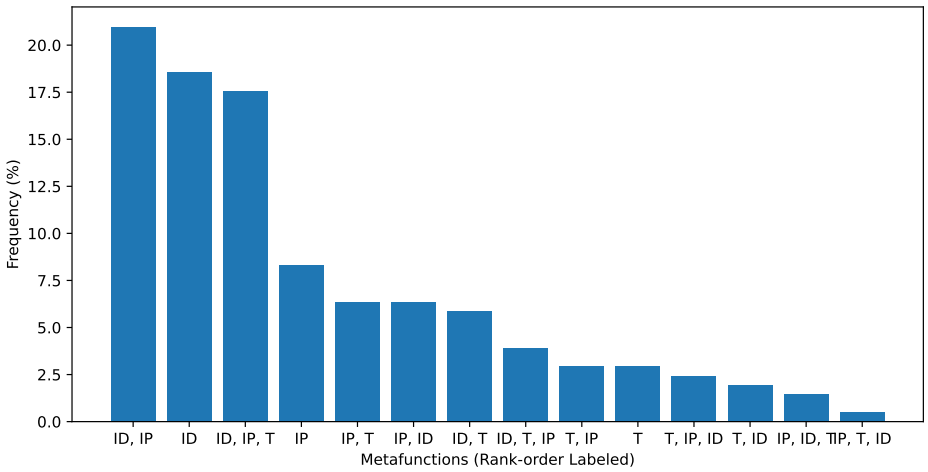

Figure 1: Distribution of adjudicated SFL metafunction labels. ID = ideational; IP = interpersonal; T = textual.

2.2 SFL Annotation

Two bilingual annotators fluent in English and Spanish independently annotated all utterances using a finalized annotation protocol grounded in SFL. Both annotators had expertise in qualitative analysis of multilingual classroom discourse. Annotations served as the reference standard for examining how psycholinguistic features align with functional meanings in multilingual student talk. Disagreements were resolved through structured discussion, with annotators jointly revisiting interpretations and reaching consensus iteratively [15].

Each utterance was labeled for the presence of one or more SFL metafunctions to reflect the multifunctional nature of classroom discourse. Annotators were permitted to assign multiple labels to a single utterance when meanings overlapped, consistent with prior discourse-analytic research on multilingual STEM classrooms. Table 1 provides an illustrative excerpt showing how ideational, interpersonal, and textual meanings can co-occur across short stretches of multilingual peer interaction.

Unit of analysis. The unit of analysis was a single speaker turn, defined as one sentence or short utterance. Utterances are understood as socially and interactionally situated units of meaning-making [1]. Multi-clause turns were annotated as a whole when they reflected a unified communicative intent.

Rank-order labeling. When multiple metafunctions were present, annotators assigned all relevant labels and ranked them by prominence, with the primary communicative function listed first. For example, the utterance *¿Qué pasó? ¿Qué vieron? [What happened? What did you see?]* was labeled as primarily ideational, reflecting the solicitation of experiential information, and secondarily interpersonal, reflecting engagement with peers and the initiation of joint sense-making. Annotators relied on expert judgment rather than fine-grained grammatical segmentation.

Annotations revealed substantial multifunctionality across the corpus. As shown in Figure 1, utterances expressing more than one metafunction were more common than single-function realizations, with ideational–interpersonal combinations occurring most frequently, followed by singular ideational turns and three-way metafunctional combinations. Purely interpersonal or purely textual utterances were comparatively rare.

2.3 Label Reliability

Inter-rater reliability was assessed using Cohen's kappa for each metafunction annotation, as seen in Table 2. Agreement was substantial for ideational meaning ($\kappa = 0.551$), high for interpersonal meaning ($\kappa = 0.693$), and moderate for textual meaning ($\kappa = 0.444$), following conventional interpretation thresholds [14]. Overall micro-averaged agreement was $\kappa = 0.595$, with an exact set-level agreement of 0.620. All disagreements were adjudicated by the primary author to produce a single gold-standard label set used in subsequent analyses.

2.4 LIWC Feature Extraction

Each utterance was analyzed using LIWC-22. LIWC produces a broad set of variables capturing summary dimensions, psychological processes, affective and social language, and function word use. Most variables are calculated as proportions of total words, while summary variables are standardized composite scores derived from underlying categories [27]. The psychometric foundations of LIWC-22 are documented in the LIWC-22 manual [4], and the English LIWC-22 and Spanish ES-LIWC2007 variables in this study are listed in the Appendix.

Because the unit of analysis is a short speaker turn, many LIWC variables were sparse or zero-inflated at the utterance level. To reduce noise associated with short texts and enable

Table 1: Excerpt of multilingual small-group science discourse with SFL metafunction labels.

Speaker	Utterance	Meta-functions
Yesenia	¿Qué pasó? ¿Qué vieron? *[What happened? What did you see?]*	Ideational, Interpersonal
Trevor	Las personas de ... o los científicos ... *[The people ... or the scientists ...]*	Ideational
Yesenia	Ajá, ¿qué están haciendo los científicos? *[Uh-huh, what are the scientists doing?]*	Ideational
Trevor	Matando a las ... um ... urchins *[Killing the ... um ... urchins]*	Ideational, Interpersonal
Yesenia	¿Erizos? *[Urchins?]*	Ideational, Interpersonal
Trevor	Erizos, yeah *[Urchins, yeah]*	Ideational
Yesenia	¿Están matando a los erizos porque qué pasó con los erizos? *[Are they killing the urchins because what happened with the urchins?]*	Ideational, Textual
Trevor	Se comen la kelpo *[They eat the kelp.]*	Ideational
Yesenia	Entonces, ¿hay muchos erizos o pocos erizos? *[So, are there many urchins or few urchins?]*	Ideational
Alonso	Muchos *[Many.]*	Ideational, Textual

Table 2: Human–human agreement for metafunction labels per utterance.

Metafunction	κ
Ideational	0.551
Interpersonal	0.693
Textual	0.444
Overall micro	0.595
Exact set-level agreement	0.620

interpretable comparison across features, LIWC variables appearing in fewer than 5% of utterances were excluded from analysis. The remaining variables were then carried forward to the correlation analyses described below, where 52 exceeded the median absolute correlation threshold for at least one metafunction and were retained for reporting.

2.5 Statistical Analysis

To examine how LIWC features correlate with functional discourse, we conducted three complementary analyses. Together, these analyses examine empirical patterns, theory-driven expectations, and robustness across language conditions.

Analysis 1: Empirical alignment between LIWC features and SFL metafunctions. First, we examined associations between LIWC variables and adjudicated SFL metafunction labels using correlation analysis. The ideational and interpersonal metafunctions were represented as binary indicators indicating whether they were present in a given utterance, which allowed both metafunctions to co-occur within the same turn.

Spearman rank-order correlations were computed between each LIWC variable and each metafunction indicator. Spearman correlation was selected to accommodate non-normally distributed variables and the ordinal nature of feature ranks. Given the exploratory goal of identifying interpretable alignment patterns, we focused subsequent analyses on LIWC variables whose correlations exceeded the median absolute correlation value for each metafunction. This thresholding strategy was used to distinguish features showing relatively stronger alignment with expert annotations from those exhibiting weak or inconsistent associations, without assuming statistical significance in a small corpus. Correlation results were visualized using heatmaps to facilitate comparison across metafunctions and LIWC feature groups.

Analysis 2: Theory-driven alignment with SFL lexicogrammatical resources. As a second line of analysis, we adopt a theory-driven approach to evaluate whether psycholinguistic features derived from LIWC align with lexicogrammatical resources theorized in SFL. Whereas Analysis 1 examined empirical associations between LIWC variables and annotated metafunctions, Analysis 2 tests whether LIWC variables hypothesized to index ideational or interpersonal meaning *a priori* exhibit differentiated alignment with these metafunctions in the annotated transcript.

Based on SFL discourse semantics and genre theory [18, 11], we specified LIWC variables expected to align with ideational and interpersonal meaning, respectively (see the Appendix). Variables were selected as surface linguistic indicators plausibly associated with different meaning orientations: ideational predictors index elaboration and references to processes or phenomena (e.g., *WC*, *WPS*, *Perception*, *space*, *differ*); interpersonal predictors index stance, social positioning, and participant orientation (e.g., *ppron*, *you*, *Social*, *Clout*). Spearman correlations were computed between each LIWC variable and metafunction.

Analysis 3: Language and Dictionary-Based Error Analysis. As a third analysis, we examined the robustness of LIWC–SFL alignment across language conditions while accounting for constraints of LIWC dictionary design. Currently, LIWC-22 supports analysis using only one language-specific dictionary at a time[1]. Accordingly, we conducted two parallel analyses across all three transcript conditions: (a) original translanguaged utterances, (b) English-only versions, and (c) Spanish-only versions using the LIWC-22 English dictionary and the internal ES-LIWC2007 Spanish dictionary.

Within each dictionary, LIWC features were computed sep-

[1]We confirmed this constraint with the creators of LIWC-22.

arately for all transcript conditions and related to the same expert-annotated SFL metafunction labels. Following prior LIWC cross-language methodology, analyses focused on within-dictionary patterns and not on direct comparisons across languages [24]. Although the English and Spanish LIWC dictionaries share category structures, subtle differences in linguistic realization and coverage exist (e.g., subject pronoun omission in Spanish), motivating a within-language analytic approach.

Together, the three analyses serve complementary purposes. Analysis 1 establishes empirical alignment patterns between LIWC features and expert interpretations of meaning. Analysis 2 evaluates whether these patterns align with theory-driven expectations about how meaning is realized lexicogrammatically. Analysis 3 probes the stability of these alignments under common preprocessing and language-restriction decisions in EDM. This design supports both exploratory and theory-driven analysis.

3. RESULTS

3.1 Analysis 1: LIWC Predictors and SFL Metafunctions

We examined associations between LIWC variables and expert-annotated SFL metafunctions using Spearman rank-order correlations. Ideational and interpersonal metafunctions were represented as binary indicators denoting whether each meaning was present in a given utterance. Fifty-two LIWC variables exceeded the median correlation threshold for at least one metafunction and were retained for analysis. For interpretability, variables were grouped post hoc into six LIWC-informed categories. Separate heatmaps were generated for each category to visualize alignment patterns with ideational and interpersonal meanings.

3.1.1 Summary Variables

Figure 4 presents correlations between SFL metafunctions and LIWC summary variables. Ideational meaning showed the strongest alignment within this category, with positive correlations for word count (WC; $\rho = 0.28$), words per sentence (WPS; $\rho = 0.28$), and BigWords ($\rho = 0.22$). Clout was also positively associated with ideational meaning ($\rho = 0.16$), while dictionary coverage (Dic) showed a negative association ($\rho = -0.18$).

Interpersonal meaning exhibited weaker and more selective alignment with summary variables. Clout showed a positive correlation ($\rho = 0.19$), while Analytic was negatively associated ($\rho = -0.23$). Other summary variables showed correlations close to zero.

3.1.2 Linguistic Dimensions

Figure 2 shows correlations for LIWC linguistic dimension variables. Interpersonal meaning exhibited the clearest alignment in this category, with the strongest associations observed for personal pronouns (ppron; $\rho = 0.28$), pronouns overall ($\rho = 0.21$), first-person singular *I* ($\rho = 0.20$), and second-person *you* ($\rho = 0.16$).

Ideational meaning showed more moderate and distributed associations across linguistic dimensions, including positive correlations for ipron ($\rho = 0.17$), determiners ($\rho = 0.17$), prepositions ($\rho = 0.17$), auxiliary verbs ($\rho = 0.15$), and conjunctions ($\rho = 0.14$). The aggregate Linguistic category was negatively associated with ideational meaning ($\rho = -0.23$).

3.1.3 Psychological Processes

Figure 3 presents correlations between metafunctions and LIWC psychological process variables. Ideational meaning showed its strongest association with *differ* ($\rho = 0.18$), followed by affiliation ($\rho = 0.11$) and cause-related language ($\rho = 0.08$). Other cognitive process categories showed weak or near-zero associations.

Interpersonal meaning exhibited modest positive correlations for discrepant language (*discrep*; $\rho = 0.12$), cognition ($\rho = 0.10$), and cognitive processing more broadly ($\rho = 0.09$), while *differ* showed a small negative association ($\rho = -0.06$).

3.1.4 Affect and Social Processes

Figure 5 shows correlations for LIWC affective and social process variables. Interpersonal meaning aligned most clearly with this category, with positive correlations for Social ($\rho = 0.22$), socrefs ($\rho = 0.18$), socbehav ($\rho = 0.12$), and assent ($\rho = 0.15$).

Ideational meaning showed weak associations with most affective and social variables, with the exception of a strong negative correlation with assent ($\rho = -0.35$), indicating that brief agreement tokens were unlikely to be annotated as ideational.

3.1.5 Perception and Time Orientation

Figure 6 presents correlations for LIWC perception and time orientation variables. Ideational meaning showed consistent positive alignment with perception ($\rho = 0.22$), spatial language ($\rho = 0.19$), motion ($\rho = 0.13$), and temporal focus, including past ($\rho = 0.12$) and future ($\rho = 0.12$) orientation.

Interpersonal meaning showed weak associations across this category, with correlations near zero for most variables and small positive correlations for present ($\rho = 0.08$) and future focus ($\rho = 0.05$).

3.1.6 Conversational and Punctuation Features

Figure 7 shows correlations for LIWC conversational and punctuation variables. Ideational meaning exhibited consistent negative correlations with conversational markers, including Conversation ($\rho = -0.29$), nonfluencies ($\rho = -0.18$), overall punctuation density (AllPunc; $\rho = -0.16$), and comma use ($\rho = -0.13$). In contrast, Apostro ($\rho = 0.12$) and OtherP ($\rho = 0.11$) showed small positive associations.

Interpersonal meaning showed weak and inconsistent alignment across conversational and punctuation features, with correlations ranging from $\rho = -0.03$ to 0.10.

3.2 Analysis 2: LIWC Alignment with SFL Lexicogrammatical Resources

We examined whether LIWC variables predicted *a priori* to align with ideational or interpersonal lexicogram-

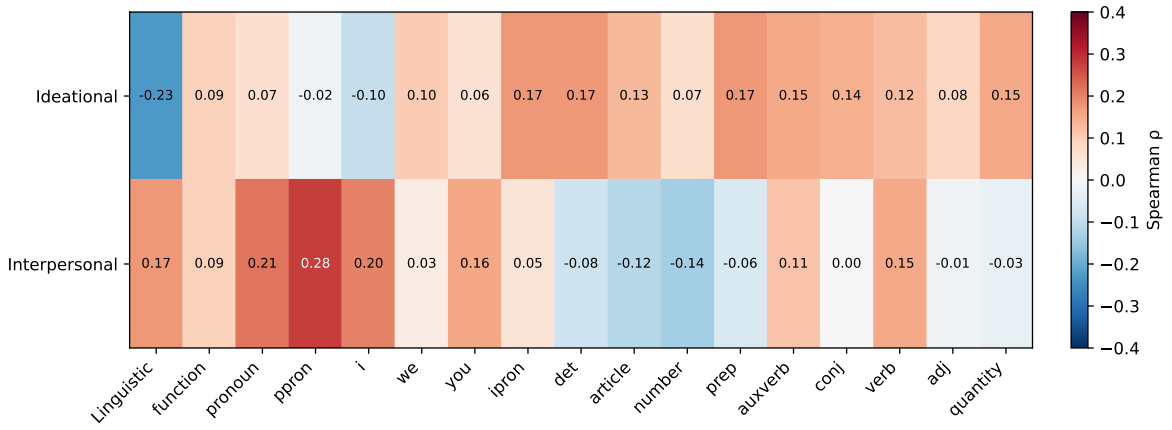

Figure 2: Spearman correlations between SFL metafunctions and LIWC Linguistic Dimensions.

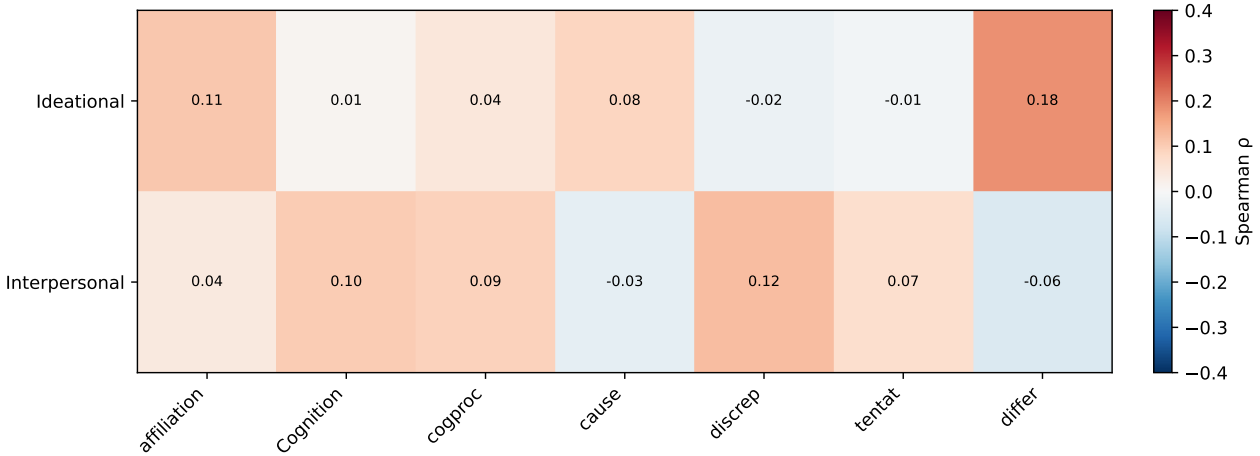

Figure 3: Spearman correlations between SFL metafunctions and LIWC Psychological Processes.

matical resources exhibit empirical alignment with expert-annotated metafunctions. For each theoretically-specified LIWC variable, Spearman correlations were computed with both ideational and interpersonal indicators to give direct comparison across metafunctions.

Ideational-predicted LIWC variables. Figure 8 shows Spearman correlations for LIWC variables hypothesized to index ideational meaning, displayed for both ideational and interpersonal metafunctions. Measures of elaboration showed the strongest positive correlations with ideational meaning, including word count (WC; $\rho = 0.28$) and words per sentence (WPS; $\rho = 0.28$). Perceptual and experiential variables also showed positive correlations with ideational meaning, including *Perception* ($\rho = 0.22$), *space* ($\rho = 0.19$), *motion* ($\rho = 0.13$), and *differ* ($\rho = 0.18$).

Across these variables, correlations with interpersonal meaning were weaker or near zero. LIWC cognitive process variables (e.g., *cogproc*, *cause*, *insight*, *discrep*) showed weak or inconsistent correlations with ideational meaning (all $\rho < 0.10$).

Interpersonal-predicted LIWC variables. Figure 9 presents correlations for LIWC variables hypothesized to reflect interpersonal meaning for ideational and interpersonal metafunctions. Pronoun-based features showed the strongest positive correlations with interpersonal meaning, including personal pronouns (*ppron*; $\rho = 0.28$), first-person singular (*I*; $\rho = 0.20$), and second-person address (*you*; $\rho = 0.16$). Social reference variables also showed positive correlations with interpersonal meaning, including *Social* ($\rho = 0.22$), *socrefs* ($\rho = 0.18$), *affiliation* ($\rho = 0.11$), and *Clout* ($\rho = 0.19$).

The same variables showed weaker correlations with ideational meaning, including *ppron* ($\rho = -0.02$), *I* ($\rho = -0.10$), *you* ($\rho = 0.06$), *Social* ($\rho = 0.13$), *socrefs* ($\rho = 0.13$), *affiliation* ($\rho = 0.11$), and *Clout* ($\rho = 0.16$).

Affective and modality-related variables showed comparatively weaker correlations with interpersonal meaning, in-

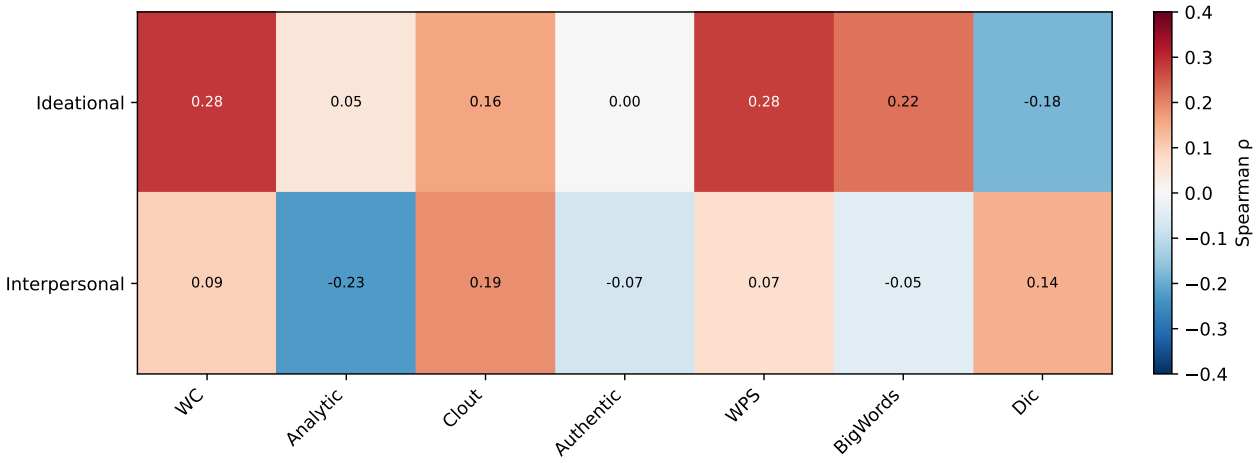

Figure 4: Spearman correlations between SFL metafunctions and LIWC Summary Variables.

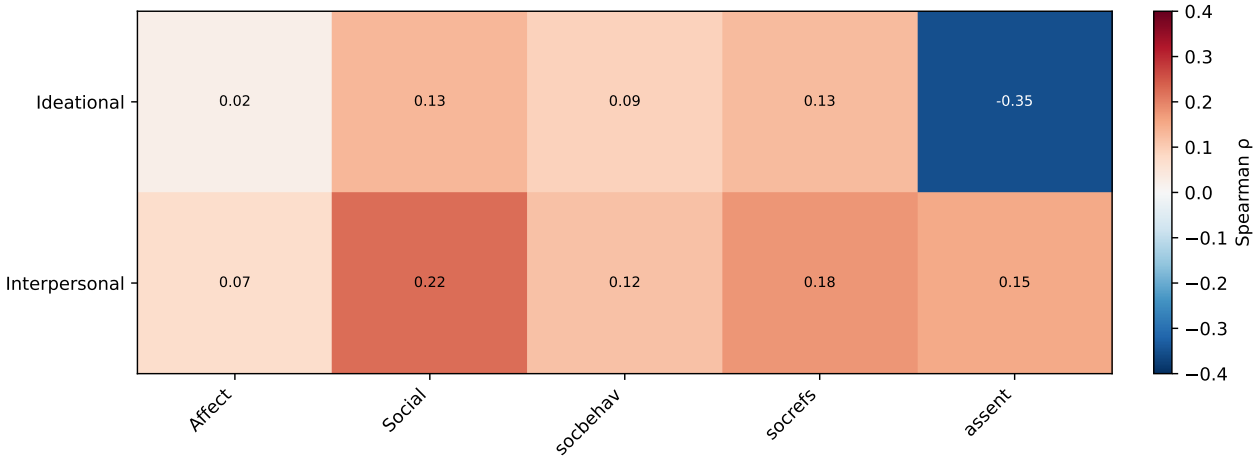

Figure 5: Spearman correlations between SFL metafunctions and LIWC Affective and Social Processes.

cluding *Affect* ($\rho = 0.07$), *tentat* ($\rho = 0.07$), and *negate* ($\rho = 0.06$).

The patterns observed in Analysis 2 support the construct validity of the theory-driven mapping. Variables predicted to align with ideational and interpersonal meaning showed differentiated correlations across metafunctions. This divergence would not be expected if LIWC features reflected only generic verbosity or affect, but instead indicates theoretical alignment with dimensions of meaning. At the same time, the weak and inconsistent performance of cognitive-process variables highlights limits in how these categories map onto interactional reasoning in multilingual classroom discourse.

3.3 Analysis 3: Language, Translation, and Error Analysis

LIWC outputs varied systematically across the original translanguaged, English-only, and Spanish-only transcript conditions, with distinct patterns depending on dictionary choice (Tables 3 and 4).

Using the LIWC-22 English dictionary (Table 3), English-only transcripts showed higher values than the original translanguaged condition for dictionary-based and function-heavy categories, including *Dic* (89.03 vs. 71.28), *Linguistic* (79.82 vs. 62.83), and *function* (55.85 vs. 42.60). The Spanish-only condition showed pronounced reductions in these same variables (e.g., *Dic* = 16.91; *function* = 8.00), while punctuation- and segmentation-related measures increased, including *Analytic* (63.81 vs. 30.43), *Period* (26.74 vs. 12.61), and *AllPunc* (57.59 vs. 42.12). Pronoun- and verb-based categories were near zero in the Spanish-only condition under the English dictionary.

Using the ES-LIWC2007 Spanish dictionary (Table 4), Spanish-only transcripts showed higher values than the original for dictionary match and grammatical categories, including *Dic* (80.70 vs. 28.28), *Funct* (51.59 vs. 19.02), *Verbos* (18.89 vs. 5.36), and *TotPron* (12.72 vs. 4.48). English-only transcripts showed consistently lower values across these same categories. Punctuation-related variables (*AllPunc*,

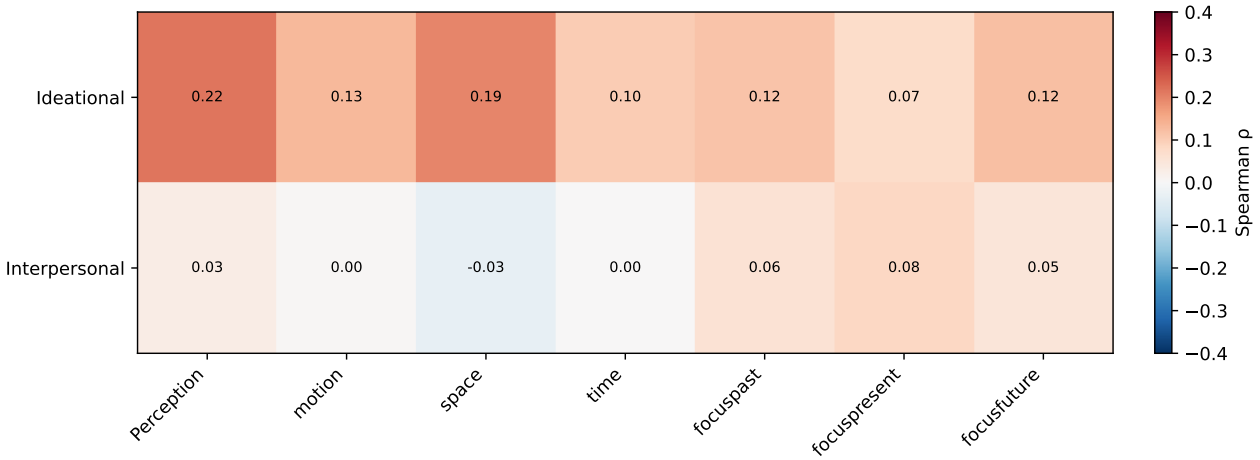

Figure 6: Spearman correlations between SFL metafunctions and LIWC Perception and Time Orientation variables.

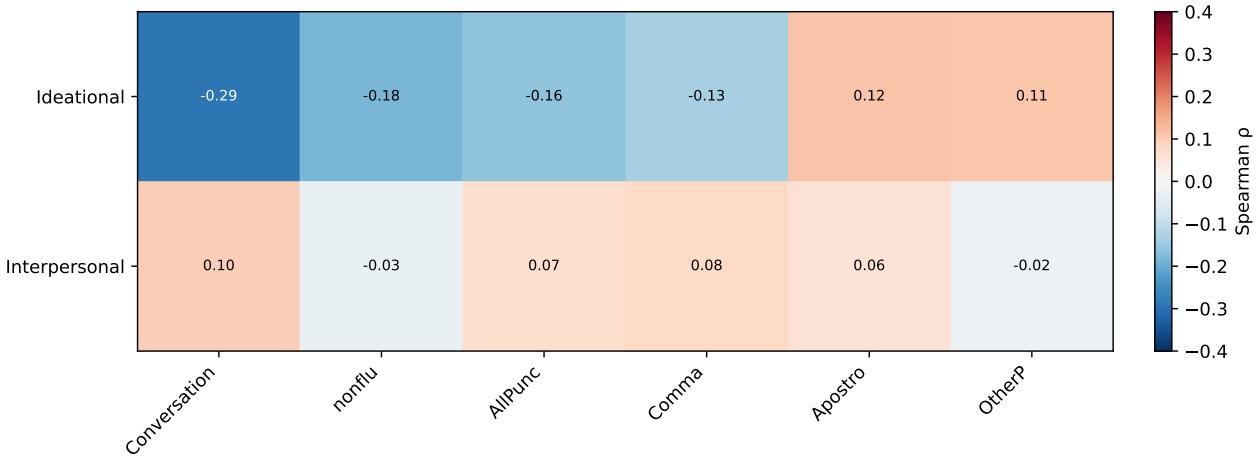

Figure 7: Spearman correlations between SFL metafunctions and LIWC Conversational and Punctuation variables.

Period, *QMark*) showed higher values in the Spanish-only condition across both dictionaries.

Across both dictionary settings, the original translanguaged condition generally fell between the English-only and Spanish-only conditions on most variables (Tables 3–4), indicating that language restriction and translation systematically reshaped LIWC feature distributions.

4. DISCUSSION

This study uses SFL to address a central challenge in EDM: how to improve interpretation of automated linguistic measures in ways that are theoretically grounded and reflect authentic language use, particularly in multilingual STEM classrooms. SFL provides a principled account of how language simultaneously construes experience, enacts social relations, and organizes discourse, making it well suited for analyzing the meaning-making practices of MLLs whose linguistic resources span languages and registers.

Against this backdrop, we examined how psycholinguistic features derived from LIWC align with expert annotations of SFL metafunctions in K-12 multilingual small-group science discourse. By grounding analysis in ideational and interpersonal meaning, the results clarify which dimensions of classroom talk are captured by commonly used psycholinguistic features and which aspects may remain less visible within current analytic approaches.

Across analyses, ideational meaning showed limited alignment with LIWC features. It was most reliably associated with surface measures of elaboration, particularly *WC* and *WPS*. Although often treated as coarse indicators, these measures indexed experiential and explanatory content in collaborative science talk, suggesting that longer and more elaborated turns were more likely to involve ideational meaning. Perceptual and spatial categories (e.g., *Perception*, *space*, *motion*) also showed modest alignment, reflecting the role of observation and spatial reasoning during scientific sense-making.

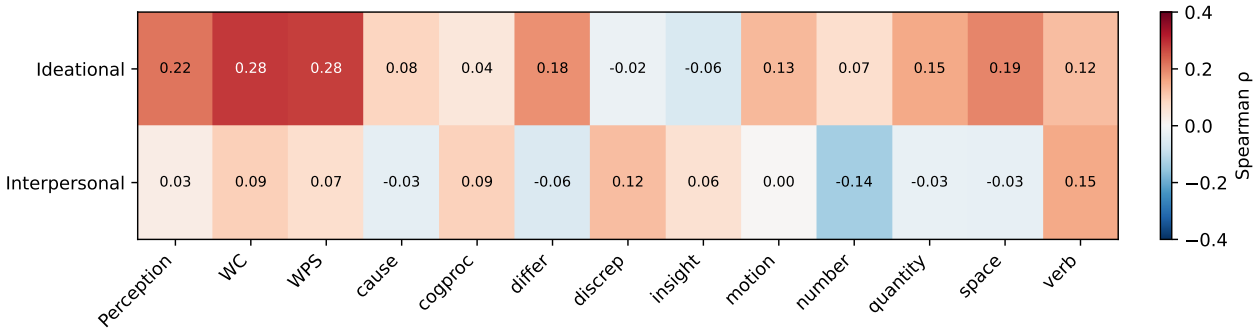

Figure 8: Spearman correlations between LIWC variables and the ideational metafunction.

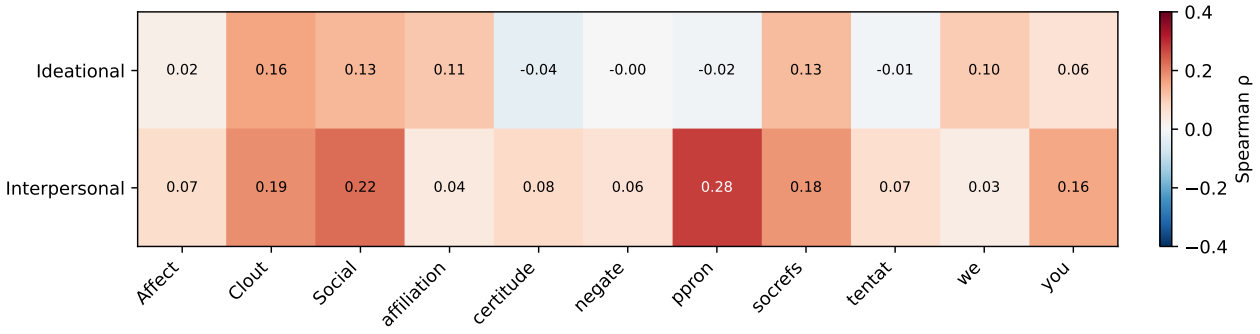

Figure 9: Spearman correlations between LIWC variables and the interpersonal metafunction.

Several LIWC variables commonly interpreted as indicators of conceptual processing showed weak or inverse relationships with ideational meaning. Higher values for *Dic* and aggregate linguistic measures (e.g., *Linguistic*, function word totals, and overall punctuation density) were negatively associated with ideational meaning. In multilingual classroom discourse, lower dictionary coverage may reflect translanguaging practices and domain-specific terminology that fall outside of monolingual dictionary inventories. As a result, utterances that are rich in disciplinary and collaborative content may register as having a lower dictionary match, despite contributing substantively to meaning-making.

In contrast, interpersonal meaning exhibited the strongest and most coherent alignment with LIWC features. Pronoun variables (e.g., *ppron*, *I*, *you*), social reference categories, and summary measures related to social positioning (e.g., *Clout*) showed the highest and most stable correlations across analyses. Agreement markers such as *assent* followed this pattern, showing a positive correlation with interpersonal meaning and a strong negative correlation with ideational meaning ($\rho = -0.35$). This pattern is most clearly reflected in Analysis 2, where variables predicted to align with ideational and interpersonal meaning showed clearly differentiated correlations across metafunctions. These patterns align with SFL accounts in which interpersonal meaning is enacted through stance-taking, evaluation, and participant positioning.

Relatedly, several LIWC categories often interpreted as indexing cognition (e.g., *cogproc*, *discrep*) aligned more strongly with interpersonal than ideational meaning. In dialogic classroom interaction, these features frequently function to signal epistemic stance, hesitation, or negotiation (e.g., "I think," "I don't know," "maybe") within interaction rather than to elaborate scientific explanations. This suggests that LIWC cognitive-process variables in collaborative settings are more closely associated with interactional positioning than with experiential or explanatory content.

Findings from the third analysis sharpen these interpretations by isolating the role of language-bound dictionary assumptions. When multilingual classroom talk was constrained to a single-language representation, the prominence of LIWC features shifted systematically, with greater emphasis on grammatical density, punctuation, and segmentation, and reduced sensitivity to social (e.g., *Social*, *socrefs*, *affiliation*) and experiential (e.g., *Perception*, *space*) categories. These shifts varied by dictionary, indicating that alignment patterns are shaped by how linguistic resources are partitioned across language-specific inventories. The translanguaged condition did not behave as an aggregation of English and Spanish features, but instead exhibited a distinct distribution of LIWC features, suggesting a redistribution of meaning-making resources that is not fully preserved under single-language dictionaries. These findings empirically challenge assumptions that academic reasoning occurs primarily in English while other languages serve only social or supportive functions, showing instead that MLLs use their full linguistic repertoires for both ideational and inter-

Table 3: Mean LIWC values across transcript conditions in the LIWC-22 English dictionary.

LIWC Variable (English dict)	Original	English-only	Spanish-only
Dic	71.28	89.03	16.91
Linguistic	62.83	79.82	12.45
function	42.60	55.85	8.00
Tone	41.83	54.45	20.60
Cognition	22.85	28.01	5.87
verb	16.39	21.61	0.00
pronoun	14.50	18.17	0.50
Clout	36.11	41.26	12.43
Social	5.92	8.53	1.13
Perception	7.22	9.29	1.24
Analytic	30.43	31.25	63.81
AllPunc	42.12	40.42	57.59
Period	12.61	13.33	26.74
QMark	12.16	8.03	17.91
BigWords	10.42	10.64	17.96

Table 4: Mean LIWC values across transcript conditions in the ES-LIWC2007 Spanish dictionary.

Spanish LIWC var	English counterpart	Original	English-only	Spanish-only
Dic	Dic	28.28	12.30	80.70
Funct	function	19.02	7.78	51.59
MecCog	Cognition	7.92	0.61	26.59
Verbos	verb	5.36	1.17	18.89
TotPron	pronoun	4.48	0.37	12.72
Present	focuspresent	4.59	1.17	16.21
Relativ	relativ	2.11	0.17	10.31
Conjunc	conj	1.23	0.00	7.23
AllPunc	AllPunc	42.12	40.42	57.59
Period	Period	12.61	13.33	26.74
QMark	QMark	12.16	8.03	17.91
Sixltr	BigWords	10.42	10.64	17.96

personal meaning-making.

Taken together, these outcomes have direct implications for how psycholinguistic features are interpreted and used in EDM models of learning in multilingual classroom contexts. In particular, LIWC categories conventionally treated as indicators of cognition or reasoning consistently aligned with interpersonal meaning in dialogic interaction. As a result, models that do not account for these learner variabilities may underestimate MLLs' knowledge and shape both assessment and downstream instructional decisions in ways that do not reflect their full meaning-making practices [13]. We therefore recommend that LIWC-based approaches more explicitly account for multilinguality, and that future work compare these features with methods that are better suited to multilingual analysis (e.g., parts-of-speech tagging).

Several limitations should be considered when interpreting these findings. The dataset consists of a relatively small number of annotated utterances drawn from focal segments of authentic classroom interactions, which enables fine-grained functional analysis but limits statistical power and generalizability to broader classroom populations. Multilingual classroom discourse highlights a significant limitation of dictionary-based approaches, as translanguaging practices and multilinguality are not fully captured by language-specific resources. In addition, SFL metafunction labels were applied at the utterance level and allowed to co-occur, reflecting the multifunctional nature of classroom talk. While theoretically grounded, this design attenuates feature–label correlations and complicates interpretation. Future work should explore representations that capture temporal context, turn-taking structure, and the development of meaning across turns, including textual and thematic organization, to better align computational features with interactional processes.

5. CONCLUSION

This study uses SFL theory to examine what LIWC features capture in multilingual classroom discourse. LIWC aligns most strongly with interpersonal meaning by capturing stance and social positioning. Ideational meaning is only partially reflected through elaboration and experiential language, while cognitive-process categories show limited alignment with explanatory content. Integrating SFL with computational analysis can support more precise interpretation of linguistic features and the development of methods that capture ideational and interpersonal meaning-making in multilingual learning contexts.

6. ACKNOWLEDGMENTS

The research reported in the article was made possible (in part) by a grant from the Spencer Foundation (#202500160). The views expressed are those of the authors and do not necessarily reflect the views of the Spencer Foundation. Support for this research was provided by the Office of the Vice Chancellor for Research and Graduate Education at the University of Wisconsin-Madison with funding from the Wisconsin Alumni Research Foundation. We thank the Multilingual Learning Research Center and Institute for Diversity Science at the University of Wisconsin-Madison for supporting our work. Ananya Ganesh was supported

by a postdoctoral fellowship from the Institute of Diversity Science. We also thank Dr. Diego Román, Dr. Ruslana Westerlund, Pragati Maheshwary, Alina Guha, and EDM reviewers for their valuable feedback.

7. REFERENCES

[1] M. M. Bakhtin. *Speech Genres and Other Late Essays*. University of Texas Press, Austin, 1986.

[2] K. Barron, N. Tseng, S. Karumbaiah, and C. T. Baeza. Evaluating language models using linguistic variations in multilingual learners' writing: A teacher study. In *Proceedings of the 19th International Conference of the Learning Sciences (ICLS 2025)*, pages 1389–1393. International Society of the Learning Sciences, 2025.

[3] S. L. Blodgett, S. Barocas, H. Daumé III, and H. Wallach. Language (technology) is power: A critical survey of "bias" in NLP. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020.

[4] R. L. Boyd, A. Ashokkumar, S. Seraj, and J. W. Pennebaker. The development and psychometric properties of liwc-22. Technical report, University of Texas at Austin, Austin, TX, 2022.

[5] M. T. H. Chi. Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1):73–105, 2009.

[6] D. Demszky, J. Liu, H. C. Hill, S. Sanghi, and A. Chung. Automated feedback improves teachers' questioning quality in brick-and-mortar classrooms: Opportunities for further enhancement. *Computers & Education*, 227:105183, 2025.

[7] O. García and A. M. Y. Lin. Translanguaging in bilingual education. In O. García, A. M. Y. Lin, and S. May, editors, *Bilingual and Multilingual Education*. Springer, 2 edition, 2017.

[8] O. García and L. Wei. *Translanguaging: Language, Bilingualism and Education*. Palgrave Macmillan, London, 2014.

[9] A. Guha, S. Karumbaiah, C. Baeza, M. Castro, K. White, and D. Roman. Teacher perceptions on ai support for bi/multilingual learners: Superpowers and constraints. In *International Conference on Artificial Intelligence in Education*, pages 85–91. Springer, 2025.

[10] M. A. K. Halliday. *An Introduction to Functional Grammar*. Edward Arnold, 1994.

[11] R. Harman. *Bilingual Learners and Social Equity: Critical Approaches to Systemic Functional Linguistics*. Springer, Cham, Switzerland, 2018.

[12] P. Joshi, S. Santy, A. Budhiraja, K. Bali, and M. Choudhury. The state and fate of linguistic diversity and inclusion in the NLP world. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020.

[13] S. Karumbaiah, A. Ganesh, S. Yang, and N. Tseng. Prompt vs. supervise: A case of using language models to assess students' science explanations. In *European Conference on Technology Enhanced Learning*, pages 244–258. Springer, 2025.

[14] J. R. Landis and G. G. Koch. The measurement of observer agreement for categorical data. *Biometrics*, 33(1):159–174, 1977.

[15] A. Lane and M. L. Kent. Dialogic engagement. *The handbook of communication engagement*, pages 61–72, 2018.

[16] J. Langman and H. Hansen-Thomas. Discourse analysis in multilingual classrooms. *Language and Education*, 2017.

[17] Y. Lin, R. Yu, and N. Dowell. Liwcs the same, not the same: Gendered linguistic signals of performance and experience in online stem courses. In *Proceedings of AIED 2020*, volume 12163 of *Lecture Notes in Computer Science*. Springer, 2020.

[18] J. R. Martin. *Discourse Semantics*. Bloomsbury Academic, London, UK, 2014.

[19] J. R. Martin and D. Rose. *Working with Discourse: Meaning Beyond the Clause*. Continuum, 2 edition, 2007.

[20] N. Mercer. Sociocultural discourse analysis: Analysing classroom talk as a social mode of thinking. *Journal of Applied Linguistics*, 1(2):137–168, 2004.

[21] R. Otheguy, O. García, and W. Reid. Clarifying translanguaging and deconstructing named languages. *Applied Linguistics*, 36(3):281–307, 2015.

[22] J. W. Pennebaker. *The Secret Life of Pronouns*. Bloomsbury, 2011.

[23] J. W. Pennebaker, C. K. Chung, J. Frazee, G. M. Lavergne, and D. I. Beaver. When small words foretell academic success: The case of college admissions essays. *PLOS ONE*, 9(12):e115844, 2014.

[24] N. Ramírez-Esparza, C. K. Chung, E. Kacewicz, and J. W. Pennebaker. The psychology of word use in depression forums in english and in spanish: Testing two text analytic approaches. In *Proceedings of the International AAAI Conference on Web and Social Media*, volume 2, pages 102–108, 2021.

[25] C. P. Rosé, O. Ferschke, V. Dutt, Y. Jo, Y. Cui, R. Wang, R. Kumar, and E. Fancher. Analyzing collaborative learning processes automatically: Exploiting the advances of natural language processing. *International Journal of Artificial Intelligence in Education*, 29(3):305–346, 2019.

[26] M. J. Schleppegrell. Review of classroom discourse analysis: A functional perspective by f. christie. *Language in Society*, 33(3):443–446, 2004.

[27] Y. R. Tausczik and J. W. Pennebaker. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of Language and Social Psychology*, 29(1):24–54, 2010.

[28] F. J. Troyan, J. D. Herazo, and M. Ryshina-Pankova. Sfl pedagogies in language education: Special issue introduction. *System*, 104:102694, 2022.

APPENDIX
A. LIWC VARIABLES

Table 5: LIWC variables used in analyses. ID = ideational; IP = interpersonal; (–) = no a priori prediction specified.

Category	Abbrev.	Description / Exemplars	Predicted
Summary Variables			
Word Count	WC	Total number of words.	ID
Analytical Thinking	Analytic	Formal, logical, hierarchical language.	ID
Clout	Clout	Confidence, leadership, social status.	IP
Authenticity	Authentic	Honest, personal, genuine language.	IP
Emotional Tone	Tone	Overall emotional positivity or negativity.	IP
Words per Sentence	WPS	Average sentence length.	ID
Big Words	BigWords	Words with six or more letters.	–
Dictionary Words	Dic	Words matched by LIWC dictionary.	–
Linguistic Processes			
Linguistic Processes	Linguistic	Aggregate function word usage.	–
Function Words	function	Articles, pronouns, prepositions, auxiliaries.	–
Pronouns (total)	pronoun / TotPron	All personal and impersonal pronouns.	–
Personal Pronouns	ppron	I, we, you, they.	IP
First-person singular	I	I, me, my.	IP
First-person plural	we	we, us, our.	IP
Second-person	you	you, your.	IP
Impersonal Pronouns	ipron	it, that, those.	–
Determiners	det	this, that, these.	–
Prepositions	prep	to, with, from, entre.	–
Auxiliary Verbs	auxverb	be, have, do.	–
Conjunctions	conj / Conjunc	and, but, or, porque.	–
Negations	negate	no, not, never.	IP
Verbs	verb / Verbos	Action and relational verbs.	ID
Quantifiers	quantity / Cuantif	many, few, much.	ID
Numbers	number	Numeric digits and number words.	ID
Adjectives	adj	Descriptive modifiers.	–
Cognitive Processes			
Cognitive Processes	cogproc / MecCog	Thinking, knowing, reasoning.	ID
Insight	insight	Realization, understanding.	ID
Causation	cause	Cause–effect relations.	ID
Discrepancy	discrep	Should, would, could.	ID
Tentativeness	tentat	Maybe, perhaps.	IP
Differentiation	differ	Contrast (but, however).	ID
Certitude	certitude	Certainty and confidence markers.	IP
Affective and Social Processes			
Affect	Affect	Emotional expression (happy, sad, angry).	IP
Social Processes	Social	Social interaction language.	IP
Social References	socrefs	References to others.	IP
Social Behavior	socbehav	Helping, talking, sharing.	IP
Affiliation	affiliation	Belonging and connection.	IP
Assent	assent	Agreement markers (yes, okay).	IP
Perception and Time			
Perception	Perception / Percept	Sensory language (see, hear, feel).	ID
Motion	motion	Movement and change.	ID
Space	space / Espacio	Spatial reference.	ID
Past Focus	focuspast	Past tense orientation.	–
Present Focus	focuspresent / Present	Present tense orientation.	–
Future Focus	focusfuture	Future-oriented language.	–
Conversational and Punctuation			
Conversation	Conversation	Informal spoken discourse markers.	–
Nonfluencies	nonflu	Filled pauses and hesitations (um, uh).	–
All Punctuation	AllPunc	Overall punctuation density.	–
Comma	Comma	Comma frequency.	–
Apostrophe	Apostro	Apostrophe frequency.	–
Other Punctuation	OtherP	Punctuation other than periods and commas.	–