

# Quantifying the Learning Process of Professional English with an Importance Weighted Absorption Metric

Mo Shi\*, Baixiao Tang

School of Economics and Management, Ankang University, Ankang, China  
Email: \*shimo0204@outlook.jp

**How to cite this paper:** Shi, M., & Tang, B. X. (2026). Quantifying the Learning Process of Professional English with an Importance Weighted Absorption Metric. *Creative Education*, 17, 1046-1074. <https://doi.org/10.4236/ce.2026.176064>

**Received:** April 15, 2026

**Accepted:** June 22, 2026

**Published:** June 25, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

The current lack of dynamic evaluation methods significantly hinders professional English instruction in higher education. To quantify the learning process in root-based teaching, this study introduces a novel Importance Weighted Absorption Metric (IWAM). Utilizing a pre-post design with 80 undergraduate students, this study collected item-level importance ratings and mastery changes for singular words and multi-word phrases. The proposed IWAM integrates student-rated relevance of pre-class with post-class mastery variations to compute weighted absorption scores. By comparing IWAM with traditional frequency-based measures, the results in this study demonstrate that importance weighting provides a more sensitive evaluation of instructional impact. The analysis in this study effectively uncovers systematic acquisition differences between single words and multi-word phrases, enabling the construction of a reproducible student knowledge-absorption index. Furthermore, this study outlines an algorithmic pathway to convert survey-derived importance ratings into classroom-level evaluational metrics. The findings offer a data-driven, process-oriented evaluation framework that can be integrated into digital teaching platforms, which is expected to optimize pedagogical interventions and enhance classroom learning efficiency.

## Keywords

Professional English, Importance Weighted Absorption Metric, Dynamic Evaluation, Root-Based Teaching, Knowledge Absorption

---

## 1. Introduction

In the contemporary society of an increasingly globalized industrial and adminis-

trative landscape, the command of English for Specific Purposes (ESP) has transitioned from an auxiliary skill to a fundamental prerequisite for undergraduates in engineering and management (Hyland, 2019). As the study of Marcu (2020) indicates, within the disciplines of the specific majors in the university, domain-specific lexical precision and formulaic multi-word expressions are not merely linguistic adornments but serve as the bedrock for both high-level technical comprehension and effective professional discourse. However, while pedagogical frameworks have widely shifted toward communicative and Task-Based Language Teaching (TBLT) to simulate real-world demands, assessment methodologies have paradoxically remained static and product-oriented, and this can be considered a critical dichotomy in current educational practices (Bygate, 2016; Reid, 2015; Robinson, 2011). Such as raw frequency counts, prevalent evaluation mechanisms predominantly rely on summative, end-of-term examinations or rudimentary quantitative metrics, which offer only a static snapshot of proficiency. Consequently, these approaches obscure the dynamic trajectory of the learning process, failing to elucidate how learners differentially allocate cognitive attention to items of varying perceived relevance or how linguistic mastery evolves temporally. This lack of granular, process-oriented data renders it challenging for instructors and curriculum designers to isolate the specific instructional variables that drive measurable knowledge absorption and long-term retention.

This disconnection manifests as a persistent methodological incongruence between pedagogical aims and the limitations of available evaluation instrumentation. Including raw frequency counts, simple test-score deltas, and binary correctness rates, traditional metrics are predicated on an erroneous assumption of lexical homogeneity. By treating distinct linguistic units as statistically equivalent, the measures above may fail to capture the nuance of differential student engagement or the specific professional value of target items (Bernaisch et al., 2022; Dudău & Sava, 2021; Evans et al., 2014). Such as root-based vocabulary building or curricula integrating discrete lexical items with formulaic sequences, the limitation is particularly detrimental in complex instructional designs due to strictly quantitative approaches that overlook the fact that the diverse item types often follow divergent acquisition trajectories. To bridge this critical gap, the present study proposes and validates a dynamic, item-sensitive evaluation framework that explicitly incorporates learner-rated importance as a weighting variable to calibrate mastery gains, thereby achieving a rigorous alignment between quantitative measurement and instructional relevance.

To operationalize this theoretical framework, this study introduces the Importance Weighted Absorption Metric (IWAM), a highly reproducible quantitative index engineered to evaluate lexical acquisition at both granular (item-specific) and aggregate (classroom) levels. The IWAM is constructed upon a bipartite empirical foundation derived from longitudinal pre- and post-instructional data, while it systematically synthesizes two distinct variables: the baseline perceived relevance (learner-rated importance) of individual lexical items prior to pedagogical

ical intervention, and the corresponding empirical deltas in item-level mastery post-instruction. By mathematically integrating item-specific importance as a modulating weight against raw mastery gains, the IWAM yields a continuous, composite absorption score. This calibration mechanism effectively amplifies the statistical signal of pedagogically meaningful knowledge acquisition while actively attenuating the “noise” associated with the rote memorization of items deemed tangentially relevant by students (Chen et al., 2021; Cowan, 2014; Tyng et al., 2017; Zhao, 2020). Upon the theories above, the metric in this study is expected to transcend traditional summative evaluation to function as a highly actionable diagnostic instrument.

In this study, the empirical phase was operationalized during the fall semester of 2025 (comprising 32 instruction hours) within a specialized, root-based professional English Class. This study consisted of 80 undergraduate students drawn from the Department of Engineering Management and Engineering Cost at Ankang University. To capture the nuanced dynamics of vocabulary acquisition, item-level data were systematically harvested via structurally aligned pre- and post-intervention instruments. Through the surveys, participants documented their baseline perceived relevance of specific linguistic targets alongside their evolving self-assessed mastery, encompassing both discrete lexical items and complex formulaic sequences. The subsequent analysis is structured to mathematically synthesize the raw survey inputs into individualized, weighted absorption scores, conducting a comparative performance evaluation against traditional frequency-based benchmarks, and delineating the divergent acquisition trajectories between distinct lexical categories.

## 2. Descriptions

### 2.1. Instructional Design Specifics

To operationalize the empirical validation of the proposed diagnostic framework, this study is situated within a specialized English for Specific Purposes (ESP) curriculum. The pedagogical intervention spanned an 8-week academic period during the fall semester of 2025 (32 contact hours), involving a purposive sample of 80 senior undergraduate students (Grade 2022) majoring in Engineering Cost and Engineering Management within the School of Economics and Management at Ankang University. The instructional scaffold is anchored to the domain-specific textbook *Professional English for Engineering Management* (co-authored by Ning, X., & Wu, C.), and the curriculum is organized into 12 thematic chapters in **Table 1**, which encompass the comprehensive engineering project lifecycle. Also, referring to **Table 2**, the targeted lexical corpus was extracted from these modules, yielding a comprehensive study pool of 316 single-word items and 262 multi-word phrases.

Rather than applying a uniform quantitative quota across all chapters, the lexical sampling density for each chapter is calibrated according to the inherent pedagogical complexity and conceptual abstractness of the specific engineering do-

main. As delineated in **Table 2**, contents such as Chapter 3 (Construction Estimating and Cost Management) and Chapter 8 (Construction Project Evaluation) exhibit the highest lexical volume. This reflects the pronounced conceptual difficulty of these subjects, which are characterized by an abundance of abstract, theoretical terminologies that lack immediate accessibility in everyday contexts. Conversely, contents with lower lexical loads, such as Chapter 2 (Construction Materials and Building Structures) and Chapter 9 (Health & Safety in Construction), deal with highly concrete, tangible concepts. These chapters utilize terminologies closely associated with empirical reality, thereby requiring a comparatively streamlined target vocabulary.

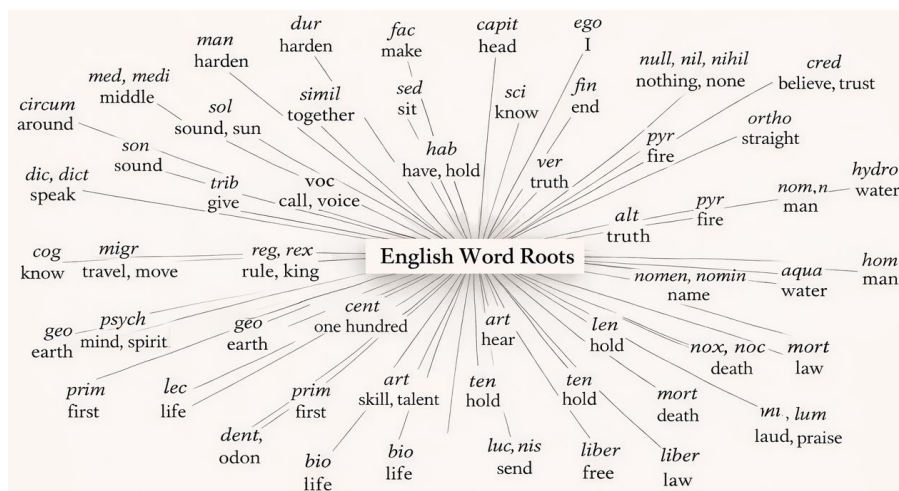
**Table 1.** Contents design of the textbook.

| Chapters   | Main Content                                     |
|------------|--|
| Chapter 1  | Construction Practice                            |
| Chapter 2  | Construction Materials and Building Structures   |
| Chapter 3  | Construction Estimating and Cost Management      |
| Chapter 4  | Construction Quality Management                  |
| Chapter 5  | Construction Scheduling                          |
| Chapter 6  | Construction Contract Management                 |
| Chapter 7  | Construction Risk Management                     |
| Chapter 8  | Construction Project Evaluation                  |
| Chapter 9  | Health & Safety in Construction                  |
| Chapter 10 | Leadership and Management in Construction        |
| Chapter 11 | Advanced Technologies in Construction Management |
| Chapter 12 | Sustainable Construction Management              |

**Table 2.** Corpus of the instruction.

| Chapters   | Word Count | Phrase Count |
|------------|------------|--------------|
| Chapter 1  | 24         | 24           |
| Chapter 2  | 33         | 7            |
| Chapter 3  | 58         | 39           |
| Chapter 4  | 23         | 8            |
| Chapter 5  | 24         | 24           |
| Chapter 6  | 17         | 17           |
| Chapter 7  | 21         | 21           |
| Chapter 8  | 41         | 41           |
| Chapter 9  | 14         | 14           |
| Chapter 10 | 27         | 27           |
| Chapter 11 | 17         | 17           |
| Chapter 12 | 17         | 23           |

Furthermore, the final compilation of the targeted items adhered to four pre-established inclusion criteria to ensure strict reproducibility and methodological rigor. First, as shown in **Table 1**, domain salience required that items embody core technical or managerial operations specific to the chapter's lifecycle stage. Second, morphological systematization dictated that single-word items possess productive Latin or Greek etymological roots (e.g., *struct-*, *fac-*, *spec-*), facilitating explicit structural deconstruction during root-based instruction. Third, such as FIDIC templates, syntactic and pragmatic Standardization restricted multi-word phrases to standard legal, contractual, or operational collocations ubiquitous in international engineering protocols. Fourth, expert panel validation was conducted to guarantee academic fidelity. The provisional pool of 316 words and 262 phrases in the instruction across this study was reviewed and verified by an independent ESP instructor team from the School of Economics and Management at Ankang University to confirm lexical appropriateness, domain accuracy, and curricular alignment.



**Figure 1.** English word roots in the lexical structure of the English language.

Within the highly contextualized matrix of professional English, the primary instructional intervention in this study is operationalized through a root-anchored pedagogical paradigm, as shown in **Figure 1**. Diverging from traditional methodologies that rely on the rote memorization of isolated vocabulary, this framework systematically orchestrated lexical acquisition around precise morphological deconstruction, including foundational roots, affixes, and derivational matrices (Beavers et al., 2017; Jassem, 2012; Turmezei, 2012). Although the linguistic nature of engineering management discourse is characterized by a high density of complex, polysyllabic technical terminology, the terms possess inherent morphological transparency once their structural components are analytically elucidated. In the pedagogical practice of this study, learners were explicitly trained to decode the internal semantic logic of specialized jargon, enabling them to systematically trace word families and map interconnected lexical networks

across the twelve thematic modules. Especially, the findings in this study indicate that this analytical approach transcended discrete word boundaries, because it actively cultivated the metalinguistic awareness of students, equipping them to logically decipher the compositional semantics of complex, multi-word formulaic sequences prevalent in engineering contexts.

In the instruction of this study, each thematic unit was operationalized through a multidimensional progression, synthesizing macro-level discourse comprehension with micro-level syntactic and lexical analysis pedagogically. Also, the targeted linguistic corpus is designed with a bifurcated typology, comprising both discrete, high-frequency technical terminology (single-word items: word) and specialized formulaic sequences (multi-word collocations: phrase). Through the subsequent application of the IWAM framework in this study, the dual-focus curriculum establishes the prerequisite structural conditions needed to empirically observe and quantitatively evaluate whether these morphologically distinct lexical forms exhibit divergent learning trajectories and differential absorption rates throughout the pedagogical intervention.

Including contextual immersion, guided morphological deconstruction, and discourse-level application, the pedagogical delivery in this study adhered to a tripartite cognitive progression to effectively operationalize this dual-focus curriculum. The design of the instruction in this study emphasizes that rather than relying on isolated memorization, students engaged in spaced retrieval practice and comparative semantic analysis to actively construct interconnected lexical networks. To longitudinally capture the dynamic trajectory of lexical absorption within this controlled setting, structured pre- and post-intervention instruments (via the Questionnaire Star digital platform) are deployed. Also, the exact same target lexical items (words & phrases) were evaluated in both the pre- and post-test instruments to guarantee longitudinal measurement invariance.

1) Perceived Importance Survey:

Administered during the pre-class phase, students responded to the explicit prompt: "Please rate the perceived importance of this lexical item to your future professional practice and academic development in engineering management." Perceptions were captured using a six-point interval scale with response anchors and corresponding numerical codings defined as follows: 0% (Completely Irrelevant, coded as 0.0), 20% (Low Relevance, coded as 0.2), 40% (Somewhat Relevant, coded as 0.4), 60% (Moderately Important, coded as 0.6), 80% (Highly Important, coded as 0.8), and 100% (Extremely Critical, coded as 1.0).

2) Linguistic Mastery Survey:

Administered both pre- and post-class, students self-assessed their knowledge depth on a 4-point ordinal mastery scale adapted from the Vocabulary Knowledge Scale (VKS), and the explicit response anchors and echelons are illustrated in **Table 3**.

Synthesize the results within perceived importance survey and linguistic

mastery survey, this study systematically recorded students' baseline perceived relevance of the targeted items prior to instruction, alongside their subsequent self-assessed mastery gains. Based on the instructional design in this study, the synergistic integration of content-based exposition, root-anchored instruction, and synchronized psychometric evaluation established a highly valid and methodologically reproducible matrix for quantifying professional English acquisition.

**Table 3.** Specifications of vocabulary knowledge scale (VKS).

| Level Score | Description   |
|-------------|---|
| A 4         | I have fully mastered this item and can deploy it productively in professional engineering discourse and project documentation. |
| B 3         | I understand the professional definition of this item and can interpret it accurately in a reading context.                     |
| C 2         | I have seen this item but do not know its definition or professional meaning.   |
| D 1         | I have never encountered this word/phrase before.   |

## 2.2. Design of the Scoring Workflow

To systematically quantify the dynamic knowledge-absorption process, the Importance Weighted Absorption Metric (IWAM) integrates cohort-level perceived importance with individual-level longitudinal mastery gains. Across item-level calculation, student-level aggregation, and class-level synthesis, the scoring workflow progresses sequentially through three distinct analytical tiers.

### 1) Item-Level Importance Weighting ( $\omega_i$ )

The baseline perceived importance of each lexical item is established at the cohort level for item-level importance weighting  $\omega_i$  before instruction. Based on the calculation method in Equation (1), let  $x_k$  represent the coded importance level assigned to category  $k$  ( $k = 1, 2, \dots, 6$ , corresponding to coded percentile values of 0.0, 0.2, 0.4, 0.6, 0.8, 1.0). For a specific lexical item  $i$ , the continuous importance weight  $\omega_i$  is computed as a weighted average across the entire student sample in Equation (1) below:

$$\omega_i = \frac{\sum_{k=1}^6 f_{ik} \cdot x_k}{\sum_{k=1}^6 f_{ik}} \quad (1)$$

where  $f_{ik}$  denotes the absolute frequency (the number of students) selecting importance category  $k$  for item  $i$ .

### 2) Individual Student Mastery Gain ( $\Delta M_{s,i}$ )

For the student and the lexical item  $i$ , the raw mastery change is defined as the empirical delta between post-instructional mastery ( $M_{s,i}^{post}$ ) and pre-instructional baseline mastery ( $M_{s,i}^{pre}$ ) in Equation (2) below:

$$\Delta M_{s,i} = M_{s,i}^{post} - M_{s,i}^{pre} \quad (2)$$

### 3) Student-Level Absorption Index ( $A_s^c$ )

Moreover, to evaluate the learning efficiency of an individual student within

a specific thematic chapter  $c$ , the student-level absorption index is calculated by modulating individual mastery gains with the cohort-level baseline importance weights across all  $N_c$  items allocated to that chapter in Equation (3) below:

$$A_s^c = \frac{1}{N_c} \cdot \sum_{i=1}^{N_c} (\omega_i \times \Delta M_{s,i}) \quad (3)$$

where  $N_c$  represents the total number of target lexical items (words or phrases) designated for chapter  $c$ .

#### 4) Class-Level Item Absorption Score ( $A_i$ )

Referring to Equation (4), the item-specific class absorption score averages the weighted gains across the entire student cohort to assess how effectively a single lexical item is absorbed at the aggregate classroom level:

$$A_i = \frac{1}{S} \cdot \sum_{s=1}^S (\omega_i \times \Delta M_{s,i}) = \omega_i \times \overline{\Delta M}_i \quad (4)$$

$$\overline{\Delta M}_i = \frac{1}{S} \cdot \sum_{s=1}^S \Delta M_{s,i} \quad (5)$$

where  $\overline{\Delta M}_i$  in Equation (5) represents the mean raw mastery change of item  $i$  across the entire class.

#### 5) Aggregate Class-Level Chapter Absorption Score ( $A_{class}^c$ )

In this study, the comprehensive macro-metric representing the macro knowledge absorption for an entire chapter is computed by averaging the item-level class absorption scores within the calculations above, as shown in Equation (6) below:

$$A_{class}^c = \frac{1}{N_c} \cdot \sum_{i=1}^{N_c} A_i \quad (6)$$

By structuring the computation through the five sequential stages above, the IWAM framework translates subjective psychometric responses into actionable, multi-dimensional educational metrics. To clarify the operational mechanics of IWAM, it is critical to state that while baseline importance ratings provide a stable cognitive weight, the individual student mastery gain ( $\Delta M_{s,i}$ ) serves as the primary dynamic engine of the overall absorption score. From item-level importance weighting ( $\omega_i$ ) and individualized mastery deltas up to the comprehensive Aggregate Class-Level Chapter Absorption Score ( $A_{class}^c$ ), the mathematical progression with the Equations (1)-(6) above establishes a reproducible evaluative continuum.

Based on the systematic synthesis of IWAM, the scoring workflow equips instructors with a diagnostic instrument that transcends traditional, static assessments, which addresses the limitation of conventional evaluations by distinguishing between items that exhibit identical raw test gains but divergent professional relevance. By utilization of IWAM, the statistical signals of gains achieved on high-priority technical terms (e.g.,  $\omega_i = 0.90$ ) are powerfully amplified, whereas

identical gains on peripheral or professionally opaque words (e.g.,  $\omega_i = 0.20$ ) are actively attenuated as evaluation noise. The mechanism of the IWAM framework in this study ensures that the final output reflects true pedagogical absorption efficiency, thereby shifting the analytical focus from basic rote memorization to targeted cognitive acquisition.

### 3. Discussion

Following the mathematical formulation and scoring workflow established in this study, the empirical analysis requires a granular examination of how the item-level importance weighting ( $\omega_i$ ) varies across the targeted linguistic corpus, while evaluating the statistical and psychometric implications of operationalizing subjective, ordinal survey ratings into a continuous, class-level composite indicator. The weighted averaging protocol embedded within the IWAM framework encompasses psychometric practicality and statistical filtering, and the six-point interval scale strikes an optimal equilibrium between measurement sensitivity and cognitive load, providing sufficient granularity to capture perceptual variance while mitigating students' fatigue and interpretative ambiguity.

Also, because student perceptions in authentic educational ecologies are intrinsically heterogeneous (Gabaldón-Estevan, 2020; Helal et al., 2018; Saadatmand & Kumpulainen, 2012; Vinje et al., 2021), relying solely on single-point central tendency estimates is methodologically insufficient. Within the systematic scoring workflow in this study, the weighted average functions as an intrinsic statistical filter by integrating the full spectrum of the cohort's response distribution (Hosseini et al., 2014; Papadakis et al., 2010; Sáez et al., 2019; Tan et al., 2018). This protocol dilutes idiosyncratic noise and isolated extreme judgments stemming from disparate prior knowledge, effectively preventing data truncation. By preserving the relative contribution of every individual rating, this approach synthesizes dispersed evaluations into a highly stable and representational baseline metric. This validated foundation sets the stage for the subsequent comparative analysis, which delineates how pedagogical interventions differentially reshape the evaluative and acquisition trajectories of discrete words and multi-word phrases.

#### 3.1. Importance Level of Words

Following the mathematical derivation of the weighted average through Equations (1)-(6), **Figure 2** delineates the longitudinal trajectory of the quantitative values, reflecting pre-class against post-class importance levels across all twelve chapters. The resulting scatter plots in **Figure 2** consistently exhibit a robust positive correlation between the initial and final evaluations for discrete lexical items. According to pre-existing pedagogical intuition regarding term relevance, this structural alignment indicates that the students' foundational epistemic schema remained highly coherent throughout the instructional cycle. Also, a systematic upward translation of the post-class distributions relative to the pre-class baselines is dis-

tinctly observable. Upon the discussions above, the upward shift signifies that the pedagogical intervention did more than merely validate prior assumptions and also catalyzed a cognitive recalibration. This study emphasizes that the instruction significantly amplified the learners' awareness of the target vocabulary's professional utility by systematically decoding morphological roots and heavily contextualizing technical terminology within the engineering management framework. Based on the analytical results in **Figure 2**, the empirical patterns substantiate a critical theoretical premise of the IWAM framework, as perceived lexical importance is not a rigid, static attribute. Conversely, it functions as a highly malleable, learnable cognitive variable that is dynamically shaped and structurally elevated through targeted pedagogical exposure and systematic morphological analysis.

To quantitatively substantiate these visual trajectories, linear regression analyses are conducted to model the predictive relationship between initial perceptions and post-class results. As detailed in **Table 4**, the fitted regression slopes predominantly range from 0.70 to 0.85 across most modules, with moderate, positive  $y$  intercepts. This specific parameter combination elucidates the mechanics of the aforementioned upward translation mathematically, which dictates a distinct pedagogical floor-raising effect as terminology that was already deemed salient pre-class maintained its high-priority status, whereas initially marginalized or less salient items experienced a proportionately larger evaluative elevation following the instructional intervention. This linear trend is particularly pronounced in Chapters 1, 3, 4, 5, 6, 8, 9, 11, and 12, where the empirical scatter points tightly conform to the positive regression vectors. From a theoretical standpoint, the regression metrics in **Figure 2** systematically confirm that content-based professional English instruction does not deconstruct or overwrite students' pre-existing lexical schemata. On the other hand, it refines and functionally amplifies them, structurally anchoring the vocabulary to domain-specific semantics, authentic usage contexts, and macro-level discourse functions.

Beyond the overarching trajectory dictated by the regression vectors, an analysis of the coefficient of determination ( $R^2$ ) unveils differences regarding the semantic stability across different chapters. As shown in **Table 4**, Chapter 4 shows the strongest linear fit ( $R^2$ : 0.7799), followed by Chapters 2 ( $R^2$ : 0.7309) and 10 ( $R^2$ : 0.7090). This stability is attributable to the semantic nature across the specific chapters, while the vocabulary therein is anchored to highly tangible, concrete engineering operations. Conversely, a stark deviation from this deterministic pattern emerges in Chapter 7, which is characterized by a substantially depressed goodness-of-fit ( $R^2$ : 0.2181) and a notably flattened slope of 0.3616, which indicates that Chapter 7 displays a high degree of evaluative dispersion. The weak linear correspondence implies that students' post-class evaluations were largely decoupled from their initial pre-class baselines, which stems from the conceptual opacity inherent to the chapter's lexical repertoire from a cognitive perspective.

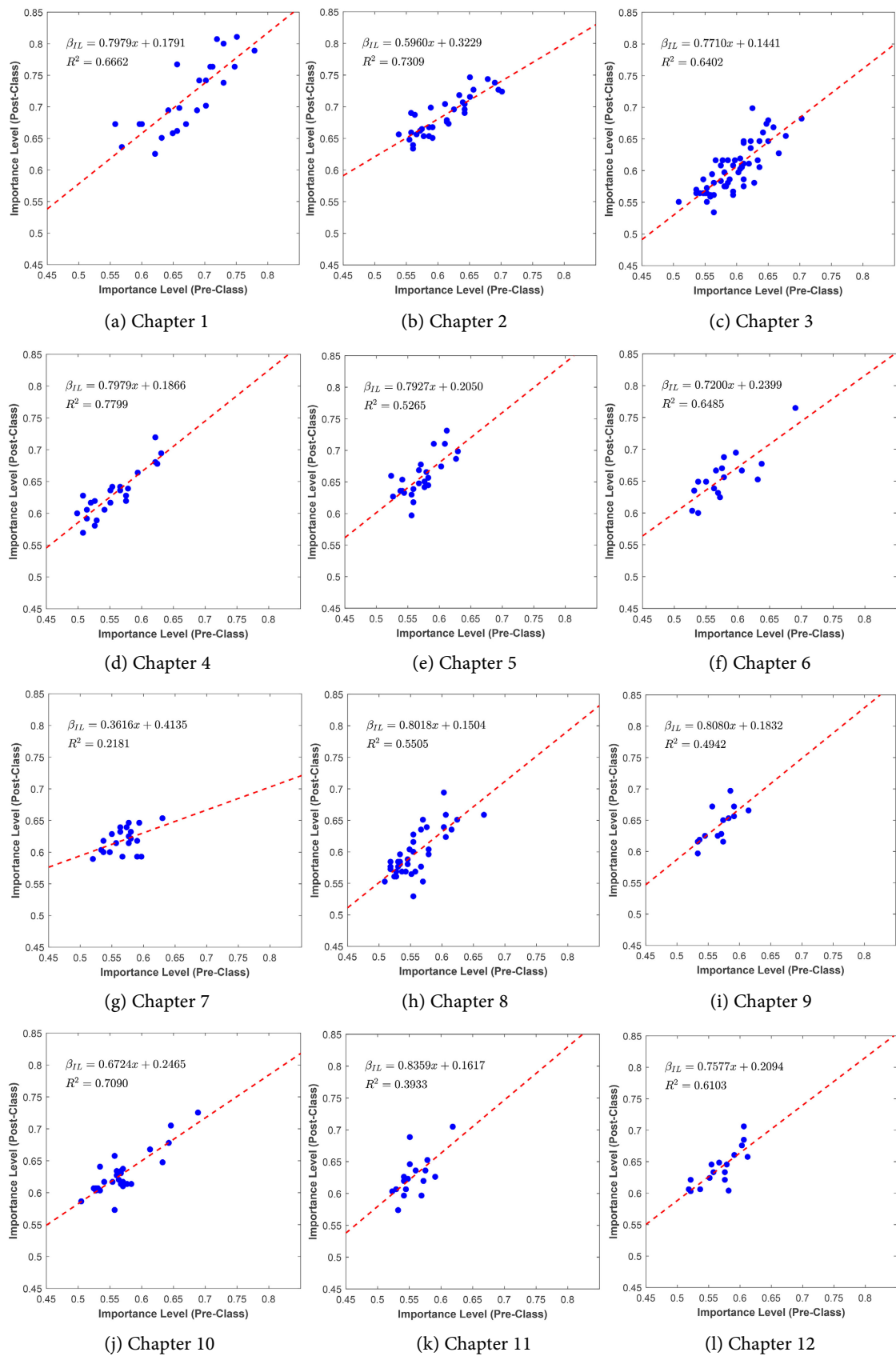


Figure 2. Evaluation of importance level (words).

**Table 4.** Linear fitting analysis of words.

| Chapters   | Slope  | Intercept | Coefficient of Determination |
|------------|--------|-----------|------------------------------|
| Chapter 1  | 0.7979 | 0.1791    | 0.6662                       |
| Chapter 2  | 0.5960 | 0.3229    | 0.7309                       |
| Chapter 3  | 0.7710 | 0.1441    | 0.6402                       |
| Chapter 4  | 0.7979 | 0.1866    | 0.7799                       |
| Chapter 5  | 0.7927 | 0.2050    | 0.5265                       |
| Chapter 6  | 0.7200 | 0.2399    | 0.6485                       |
| Chapter 7  | 0.3616 | 0.4135    | 0.2181                       |
| Chapter 8  | 0.8018 | 0.1504    | 0.5505                       |
| Chapter 9  | 0.8080 | 0.1832    | 0.4942                       |
| Chapter 10 | 0.6724 | 0.2465    | 0.7090                       |
| Chapter 11 | 0.8359 | 0.1617    | 0.3933                       |
| Chapter 12 | 0.7577 | 0.2094    | 0.6103                       |

Synthesizing the empirical insights derived from the preceding regression curves and scatter plots in **Figure 2**, the pervasive upward displacement of the evaluative trajectories underscores that the instructional intervention successfully engineered an epistemic reevaluation of the target lexis. Rather than merely transmitting definitional equivalence, the highly contextualized pedagogy compelled students to reinterpret previously underestimated terminology through professional views, fundamentally elevating their perceived utilitarian value. From a cognitive acquisition standpoint, the pedagogically induced shift is critical, as heightened perceived relevance directly catalyzes optimized attentional allocation and enhances long-term retention probabilities. Moreover, from a methodological paradigm, the empirical clustering around the chapter trends definitively validates the weighted average approach as a mechanism for filtering stochastic individual variance and capturing stable, class-level consensus.

### 3.2. Importance Level of Phrases

Shifting the analytical aspects from discrete lexical items to multi-word formulaic sequences as phrases, **Figure 3** illustrates the pre- and post-class evaluative trajectories across all the chapters. Reflecting the macroscopic statistical trends observed in the word-level analysis, the scatter plots for phrases exhibit a positive correlation alongside a systematic upward trend of post-class scores. This structural alignment confirms that students maintain a coherent, non-random baseline of pragmatic intuition even for complex linguistic units. On the other hand, a comparative analysis reveals that multi-word phrases display a significantly heightened degree of instructional sensitivity. Because the professional and semantic significance of a formulaic sequence frequently supersedes the mere literal sum of its constituent words, the pre-class utility of the phrases is often more pragmatically opaque to students. Upon the discussion above, while discrete words heavily benefit from morphological deconstruction, the perceived im-

portance of phrases is far more profoundly shaped and elevated by the instructor's discourse-level contextualization, syntactic parsing, and repeated exposure within authentic professional texts.

To quantitatively substantiate the hypothesized instructional sensitivity in this study, the regression parameters detailed in **Table 5** offer mathematical evidence. The fitted linear models across the modules yield positive slopes predominantly ranging from 0.60 to 0.91, coupled with moderate to high coefficients of determination. The parametric profile indicates that while the post-class evaluation of phrases remains structurally tethered to pre-class baselines, it undergoes significant pedagogical modulation. Due to several chapters exhibiting conspicuously high positive intercepts, the most informative metric lies in the  $y$  intercepts. The specific combination in this study dictates a pronounced compensatory evaluative lift due to a fractional slope paired with a substantial intercept. It implies that initially marginalized phrases, which possessed modest pre-class salience due to their pragmatic opacity, received a large elevation in importance post-instruction, which encapsulates the transformative function of professional English teaching pedagogically.

Beyond the overarching trend of compensatory lift, a granular examination of the coefficient of determination exposes significant chapter variations in semantic stability. As indicated in **Table 5**, Chapter 10 commands the highest explained variance ( $R^2$ : 0.8139), closely followed by Chapters 2, 3, and 12 ( $R^2$ : approx 0.75), which implies that the formulaic sequences within these chapters maintained a highly predictable evaluative trajectory. Upon the discussion with the higher coefficient of determination in **Table 5**, the results suggest that the phrases in these chapters inherently possess a higher degree of internal lexical transparency, allowing students to form consistent semantic expectations that align smoothly with subsequent instruction. In stark contrast, Chapters 6 and 7 exhibit severely depressed goodness-of-fit metrics ( $R^2$ : 0.2093 & 0.3207), accompanied by markedly flattened slopes (0.3179 & 0.3664), which indicate intense evaluative dispersion. Based on the lower coefficient of determination, the volatility is symptomatic of high thematic abstraction, emphasizing that the phrases in these chapters are largely functionally complex and pragmatically opaque, rendering them exceedingly difficult for novices to evaluate without heavy pedagogical scaffolding and deep contextual support.

While the dominant statistical paradigm across most chapters is characterized by the compensatory lift due to the slope being less than 1.0, an inspection of the regression parameters reveals compelling structural anomalies in Chapters 9 and 11. Referring to **Table 5**, Chapter 9 generates a regression slope that distinctly exceeds unity. Mathematically, a slope greater than 1.0 dictates an amplification of pre-existing variance, implying that the instructional intervention does not simply compress the gap between salient and marginal phrases, but rather magnifies the hierarchical distinctions, engendering a phenomenon of evaluative polarization. On the other hand, Chapter 11 exhibits a near-unity slope, denoting a virtually isomorphic transformation where the pre-instructional structural distances between phrase values were preserved while experiencing a proportional upward translation.

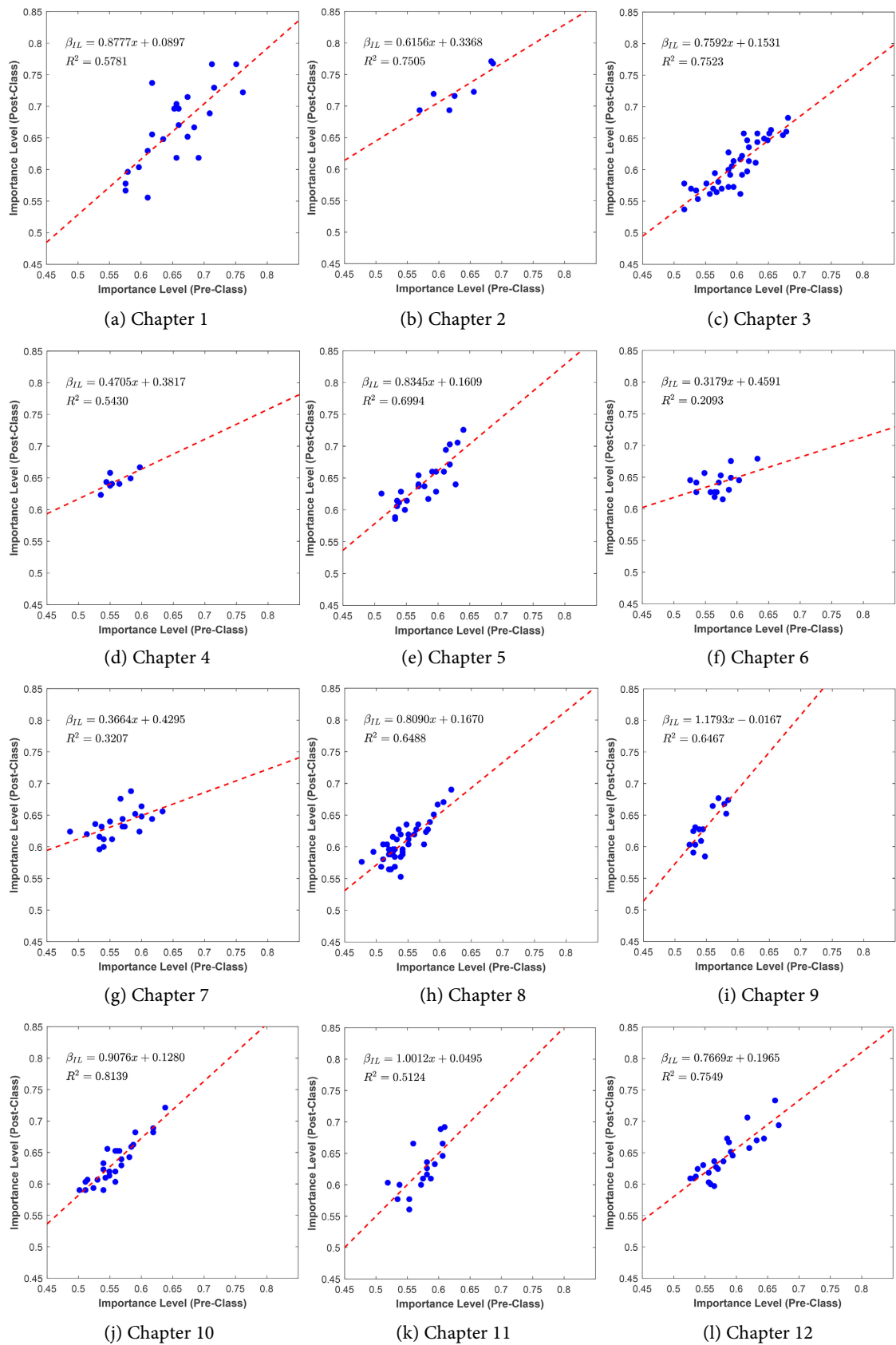


Figure 3. Evaluation of importance level (phrases).

**Table 5.** Linear fitting analysis of phrases.

| Chapters   | Slope  | Intercept | Coefficient of Determination |
|------------|--------|-----------|------------------------------|
| Chapter 1  | 0.8777 | 0.0897    | 0.5781                       |
| Chapter 2  | 0.6156 | 0.3368    | 0.7505                       |
| Chapter 3  | 0.7592 | 0.1531    | 0.7523                       |
| Chapter 4  | 0.4705 | 0.3817    | 0.5430                       |
| Chapter 5  | 0.8345 | 0.1609    | 0.6994                       |
| Chapter 6  | 0.3179 | 0.4591    | 0.2093                       |
| Chapter 7  | 0.3664 | 0.4295    | 0.3207                       |
| Chapter 8  | 0.8090 | 0.1670    | 0.6488                       |
| Chapter 9  | 1.1793 | 0.0167    | 0.6467                       |
| Chapter 10 | 0.9076 | 0.1280    | 0.8139                       |
| Chapter 11 | 1.0012 | 0.0495    | 0.5124                       |
| Chapter 12 | 0.7669 | 0.1965    | 0.7549                       |

Synthesizing the empirical evidence derived from the preceding analyses based on **Figure 3** and **Table 5**, the findings indicate that the cognitive appraisal of multi-word phrases is a highly dynamic and non-monolithic trajectory. Unlike discrete words, phrase mastery in professional English transcends simple lexical recognition, which demands advanced cognitive processing, including syntactic chunking, collocation awareness, and domain-specific discourse interpretation. Because of this inherent complexity, the recalibration of phrase importance is heavily moderated by the interplay between thematic complexity and intrinsic lexical transparency. Following the discussion above, instructional intervention functions either as a baseline consolidator on the specific textual demands or as an analytical amplifier that deliberately intensifies lexical stratification to cultivate granular evaluative judgment. Methodologically, the capacity of the weighted average to accurately capture the chapter-specific, multidimensional shifts, while maintaining sufficient stability to filter out individual stochastic noise. Based on the IWAM analysis in this study, the findings validate that phrase importance is not only mathematically measurable and acutely instruction-responsive, but also serves as a highly reliable intermediate variable.

### 3.3. Comparison between Words and Phrases

Transitioning from the isolated analyses of discrete lexical items and formulaic sequences, **Figure 4** constructs a macroscopic, full-cycle comparative synthesis between words and phrases within the overarching IWAM framework. According to both analytical results, the overarching positive association between pre- and post-class evaluative metrics confirms a shared baseline of structural continuity throughout the instructional cycle. However, a direct parametric comparison reveals that these two lexical typologies respond to pedagogical intervention with distinct cognitive trajectories. As shown in **Figure 4**, the regression model for dis-

crete words yields a steeper gradient and tighter linear fidelity, whereas multi-word phrases generate a noticeably flatter slope coupled with increased evaluative variance. Statistically and pedagogically, the findings indicate that students' value judgments regarding isolated words maintain a relatively stable cognitive anchorage from pre-class to post-class, largely preserving their initial hierarchical rankings (Bireta & Mazzei, 2016; de Graaff et al., 2008; Hyönä et al., 2002; Joshi et al., 2014; Robinson et al., 2012). Conversely, the appraisal of phrases exhibits a heightened pedagogical plasticity due to the weaker linear continuity, which signifies a more profound instructional recalibration, demonstrating that the perceived importance of complex sequences is fundamentally more dependent on classroom exposure and teacher-guided discourse analysis.

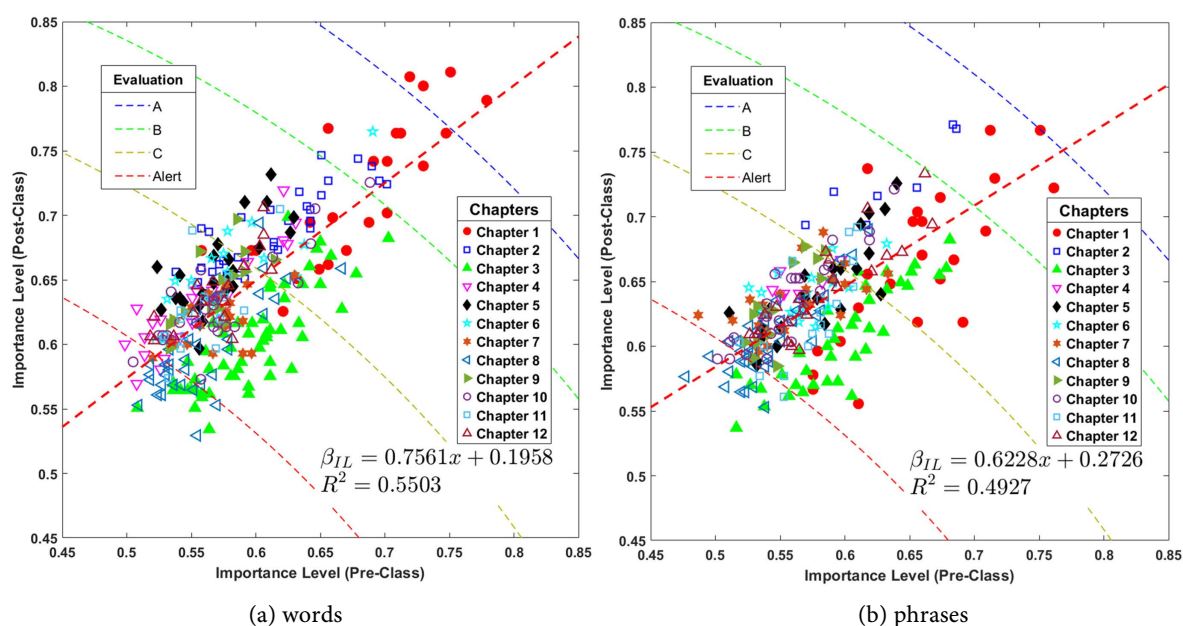


Figure 4. Full cycle comparison analysis.

Moving beyond the overarching regression metrics, a granular spatial analysis of the scatter plot distributions across the predefined evaluation zones (Alert, C, B, and A) further substantiates the cognitive dichotomy. As shown in Figure 4, both lexical typologies exhibit a pronounced central tendency, with a substantial density of observations clustering within the intermediate strata. Regardless of their morphological category, the structural convergence indicates that the majority of target items naturally gravitate toward a baseline of moderate pedagogical salience rather than occupying the extreme peripheral zones. However, a critical divergence emerges in their upward mobility post-instruction because the results in Figure 4 indicate that the discrete word dataset demonstrates a significantly broader upward dispersion, characterized by a robust subset of observations successfully penetrating the upper-tier evaluation bands (Regions B and A), which indicates that discrete vocabulary items are more readily catalyzed by instruction to achieve advanced stages of lexical internalization. In contrast, multi-word

phrase observations remain densely confined to the median cognitive strata, exhibiting a distinct “ceiling effect” that restricts their mass migration into the highest echelons of perceived importance (Bruijnzeel et al., 2017; Hammers et al., 2024; Rasmussen et al., 2001; Yitzhak et al., 2016). Including compositional complexity, collocational dependence, and high contextual variability, the findings emphasize that the restricted upward mobility strongly suggests that phrase acquisition is heavily burdened by compounding cognitive loads (Jiang et al., 2019; Shadrova, 2025; Xu & Yu, 2025). Due to this, the findings in this study note that while phrases heavily benefit from compensatory instructional lift at the bottom, their intrinsic structural opacity inherently impedes rapid, full-scale ascension into the absolute highest categories of communicative value.

To operationalize the cognitive and spatial disparities into actionable teaching strategies, the MATLAB-coded evaluation framework in Figure 5 establishes a four-tier decision-making matrix for differentiated instruction. As delineated by the thresholds in Figure 4, items falling into the Alert zone (metric < 0.75) denote persistently weak bidirectional evaluation, thereby mandating highly targeted pedagogical reinforcement to optimize learning efficiency. The C region (0.75 ≤ metric < 0.85) encapsulates a moderately competent knowledge baseline that requires systematic maintenance and progressive enhancement through contextualized simulations and structured retrieval. Moving up the continuum, the B region (0.85 ≤ metric < 0.95) signifies an upper-intermediate mastery, while the well-internalized items function as critical “bridging knowledge” rather than being treated in isolation. Pedagogically, they can be strategically and organically integrated to provide a cognitive scaffold for the acquisition of weaker items residing in the Alert and C zones. At the apex, the A region (metric ≥ 1.05) captures exceptional evaluative outcomes, serving as potent motivational anchors to stimulate student engagement while synergistically elevating the entire lexical network. In this study, the architectural advantage of the MATLAB-based design lies in its parametric flexibility, which indicates that teachers can dynamically calibrate these specific numerical boundaries based on empirical teaching experience and cohort-specific cognitive baselines, rather than imposing rigid, immutable classifications and situated instructional goals, thereby ensuring the evaluation matrix remains highly adaptive to real-world pedagogical contexts.

```
theta = linspace(0, pi/2, 200);
h_alert = plot(ax1, 0.85*cos(theta), 0.75*sin(theta), '--r', 'LineWidth', 1, 'DisplayName', 'Alert');
h_c = plot(ax1, 0.95*cos(theta), 0.85*sin(theta), '--y', 'LineWidth', 1, 'DisplayName', 'C');
h_b = plot(ax1, 1.05*cos(theta), 0.95*sin(theta), '--g', 'LineWidth', 1, 'DisplayName', 'B');
h_a = plot(ax1, 1.10*cos(theta), 1.05*sin(theta), '--b', 'LineWidth', 1, 'DisplayName', 'A');
```

**Figure 5.** Evaluation framework on MATLAB.

The comparative synthesis presented in Figure 4 elucidates that while word-level importance exhibits smooth, predictable continuity, phrase-level appraisal is intricately shaped by a complex confluence of syntactic, collocational, and discourse-level variables. Referring to Figure 4, the diminished coefficient of determination and the concentrated middle-band distribution observed for phrases

must not be construed as a lack of statistical reliability, but signify a heightened contingency on pedagogical framing. The findings in this study indicate that discrete words should be systematically leveraged to consolidate foundational terminological networks, whereas formulaic phrases necessitate explicit contextual embedding and iterative functional practice to propel students from the intermediate Alert/C baseline into the advanced B/A mastery echelons. By successfully distinguishing not only pre- and post-instructional cognitive states but also the nuanced absorption signatures of heterogeneous lexical forms, the analysis in this study profoundly validates the interpretive power of the IWAM framework, while establishing IWAM as a robust, process-oriented metric essential for the fine-grained diagnostic evaluation of classroom learning efficiency.

To further substantiate the mechanistic divergences with granular aggregate metrics, the aggregate pre-class means (0.5814 for words, 0.5802 for phrases) and post-class means (0.6424 for words, 0.6418 for phrases) exhibit remarkable statistical isomorphism, followed by the analytical results in **Table 6**. The baseline parity in this study robustly eliminates initial item-selection bias, ensuring that the ensuing variations in learning trajectories stem genuinely from instructional effects rather than inherent disparities in perceived baseline difficulty. While both lexical categories demonstrate a definitive upward trajectory, an examination of the comparative gains reveals a consistent proportional advantage for phrases (Absolute Gain: 6.16%, Relative Gain: 10.89%) over discrete words (Absolute Gain: 6.10%, Relative Gain: 10.72%), while the growth is characterized by pronounced chapter-level heterogeneity. Specifically, discrete words achieved their peak absolute gains in Chapter 5 (8.65%) and Chapter 6 (7.77%), whereas multi-word phrases experienced their most potent instructional surges in Chapter 2 (9.37%) and Chapter 4 (8.54%), juxtaposed against severe stagnation in Chapters 1 and 3 (both under 1.00%). This sharply fluctuating, chapter-dependent dispersion indicates that, unlike discrete words, which consolidate relatively evenly across standard pedagogical exposures, phrase acquisition is intensely episodic and highly contingent upon the specific thematic affordances, explicit collocational instruction, and discourse-level contextualization embedded within distinct chapters.

In this study, the quantitative evidence delineated in **Table 6** robustly elucidates that IWAM supersedes conventional, static frequency-based paradigms. The analytical results in **Table 6** indicate that the documented baseline parity across disparate lexical forms confirms that IWAM does not merely mirror superficial lexical exposure, while effectively capturing authentic, instructionally-driven cognitive recalibrations in student perception. Also, the deployment of dual gain indices precisely accommodates the inherent non-uniformity of the learning process, because absolute gain quantifies the raw magnitude of cognitive shift, and relative gain critically standardizes proportional growth against varying initial baselines, making it an indispensable parameter for cross-chapter comparative analyses. Based on the discussion above, the approach of IWAM possesses the architectural capacity to simultaneously represent the structural stability of discrete words and

the high instructional malleability of multi-word phrases, while rendering the metric theoretically rigorous, empirically highly sensitive, and exquisitely suited as a process-oriented instrument for tracking the intricate absorption trajectories in professional English pedagogy.

**Table 6.** Averaging analysis of the importance level.

| Types   | Chapters   | Importance level |            | Absolute Grain | Relative Gain |
|---------|------------|------------------|------------|----------------|---------------|
|         |            | Pre-Class        | Post-Class |                |               |
| words   | Chapter 1  | 0.6737           | 0.7167     | 4.30%          | 6.59%         |
|         | Chapter 2  | 0.6101           | 0.6865     | 7.64%          | 12.80%        |
|         | Chapter 3  | 0.5953           | 0.6031     | 0.78%          | 1.42%         |
|         | Chapter 4  | 0.5563           | 0.6304     | 7.42%          | 13.51%        |
|         | Chapter 5  | 0.5716           | 0.6581     | 8.65%          | 15.23%        |
|         | Chapter 6  | 0.5792           | 0.6570     | 7.77%          | 13.62%        |
|         | Chapter 7  | 0.5684           | 0.6191     | 5.06%          | 9.05%         |
|         | Chapter 8  | 0.5596           | 0.5991     | 3.95%          | 7.15%         |
|         | Chapter 9  | 0.5681           | 0.6422     | 7.41%          | 13.10%        |
|         | Chapter 10 | 0.5695           | 0.6294     | 5.99%          | 10.73%        |
|         | Chapter 11 | 0.5569           | 0.6272     | 7.03%          | 12.67%        |
|         | Chapter 12 | 0.5681           | 0.6398     | 7.17%          | 12.73%        |
| phrases | Chapter 1  | 0.6579           | 0.6671     | 0.93%          | 1.49%         |
|         | Chapter 2  | 0.6325           | 0.7263     | 9.37%          | 15.05%        |
|         | Chapter 3  | 0.5973           | 0.6065     | 0.92%          | 1.68%         |
|         | Chapter 4  | 0.5596           | 0.6450     | 8.54%          | 15.34%        |
|         | Chapter 5  | 0.5765           | 0.6420     | 6.55%          | 11.48%        |
|         | Chapter 6  | 0.5696           | 0.6402     | 7.05%          | 12.55%        |
|         | Chapter 7  | 0.5625           | 0.6356     | 7.31%          | 13.30%        |
|         | Chapter 8  | 0.5427           | 0.6060     | 6.34%          | 11.77%        |
|         | Chapter 9  | 0.5494           | 0.6312     | 8.18%          | 14.88%        |
|         | Chapter 10 | 0.5577           | 0.6341     | 7.65%          | 13.79%        |
|         | Chapter 11 | 0.5737           | 0.6239     | 5.02%          | 8.77%         |
|         | Chapter 12 | 0.5826           | 0.6433     | 6.07%          | 10.58%        |

Extending the aggregate gain evaluation, **Figure 6** introduces a full-cycle, chapter-wise mapping of the regression dynamics for both lexical categories under the IWAM framework, showing that a ubiquitous baseline feature across all twelve chapters is the manifestation of consistently positive regression slopes. The structural convergence indicates that the pedagogical intervention preserves the structural fidelity of baseline lexical relevance, indicating post-class importance scales proportionately with pre-class perception, amplifying rather than distorting the

students' foundational cognitive schema. However, a quantitative synthesis of the regression parameters reveals that the mean regression slope for multi-word phrases exhibits a marginal discernible superiority over that of discrete words, denoting a slightly stronger overall instructional amplification effect, and the pronounced disparity in their standard deviations mathematically corroborates the episodic learning hypothesis established earlier. While discrete words maintain a highly stable and compact evolutionary trajectory characterized by low variance, phrase acquisition is subject to profound cross-chapter volatility, as evidenced by a standard deviation nearly twice as large. Epistemologically, this distinct variance profile powerfully refutes the reductive assumption that multi-word expressions are merely scaled-up, "harder" iterations of discrete vocabulary (Arana et al., 2024; Schwartz et al., 2013; Zaninello & Birch, 2020). In contrast, it definitively categorizes them as fundamentally distinct pedagogic entities, and perceived communicative value is acutely sensitive to thematic fluctuations, explicit contextual embedding, and localized instructional emphasis.

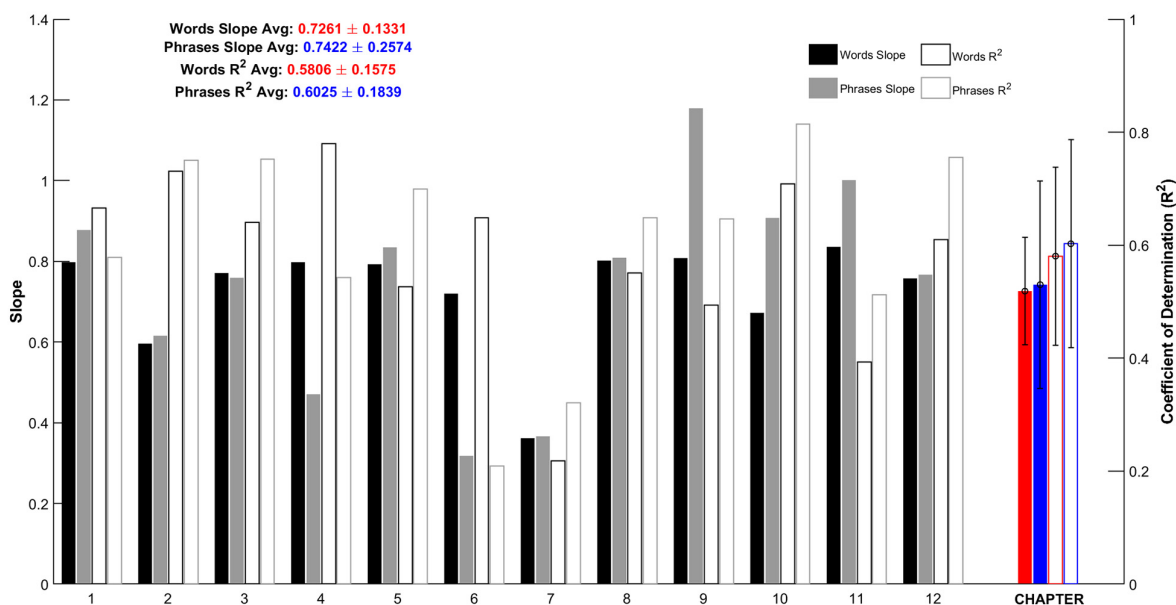


Figure 6. Comparison of increasing & fitting.

Reflecting the slope dynamics, a comparative analysis of the coefficients of determination ( $R^2$ ) further reinforces this mechanistic dichotomy. As illustrated in Figure 6, the mean  $R^2$  value for multi-word phrases marginally eclipses that of discrete words, indicating that, on a macro scale, the post-instructional importance of phrases retains a robust linear dependence on their pre-class baselines. On the other hand, mirroring the slope variance, the  $R^2$  distribution for phrases exhibits a substantially wider dispersion, which signifies acute chapter dependency as specific thematic contexts catalyze highly coherent, predictable phrase-learning trajectories, whereas other instructional settings yield diminished linear continuity. Moreover, the  $R^2$  metrics for discrete words manifest remarkable

cross-chapter stability as shown in **Figure 6**, which enduring consistency logically aligns with the inherent morphological nature of individual words, which are structurally more amenable to isolation, explicit annotation, and retention via standardized root-based instructional paradigms. Synthesizing the slope and  $R^2$  profiles, it becomes evident that while both lexical categories share a congruent overarching developmental trajectory, they diverge fundamentally in their pedagogical volatility and sensitivity to situated instructional contexts.

Based on the view through the IWAM framework, the regression dynamics encapsulated in **Figure 6** substantiate the methodological imperative of integrating explicit importance weighting with absorption-based change analysis. This study emphasizes that the evaluation relied exclusively on raw frequency counts or unweighted summative test scores, the profound mechanistic distinctions between discrete words and formulaic phrases would be critically obfuscated, as flattened by reductionist metrics into a monolithic and undifferentiated aggregate of success. In stark contrast, the multidimensional architecture of IWAM empowers teachers to decode not merely the binary occurrence of learning, but absolute instructional magnitude, cross-chapter pedagogical consistency, and the underlying structural robustness of the cognitive shifts involved. As indicated in **Figure 6**, the distinct variance signatures embedded within the slope and  $R^2$  distributions conclusively verify that the weighted importance parameter is both robustly measurable and exquisitely sensitive to instructional nuances. Also, its capacity to faithfully reflect chapter-level volatility demonstrates an inherent contextual elasticity tailored to authentic pedagogical realities.

#### 4. Summarizations

Transitioning from the quantitative trajectory of the IWAM framework to the linguistic intrinsic properties of the evaluated items, **Table 7** elucidates a striking semantico-syntactic dichotomy between the highest- and lowest-ranked vocabulary across the twelve chapters. The apex of the importance hierarchy is systematically dominated by items exhibiting high operational tangibility and domain embeddedness, as the tier is heavily populated by action-oriented verbs (e.g., managing, tracking, implementing) and entity-bound nouns (e.g., sketch, codes, audits). Within the discourse of professional engineering, the categories bear the primary informational load because they denote explicit managerial actions, concrete processes, and tangible deliverables. Their immediate mappability onto real-world construction management engineering endows them with high pragmatic utility, rendering them highly salient and pedagogically accessible to students. In stark contrast, the lowest-ranked echelon is characterized by significant conceptual distance as this tier exhibits a disproportionate prevalence of abstract nouns (e.g., taxonomies, externalities), alongside adjectives and adverbs that convey stance, degree, or theoretical categorization (e.g., iterative, inherently, euphemistically). Rather than describing direct professional interventions, these lexemes function within an evaluative, interpretive, or disciplinary-auxiliary register. Due to this, the

semantic opacity and detachment from the visible engineering workflow render them less immediately “usable” to students. In conclusion, this polarized distribution confirms that students implicitly prioritize lexical items with direct operational affordances, consistently assigning lower importance scores to conceptually dense or rhetorically specialized vocabulary that lacks immediate vocational resonance.

**Table 7.** Word items with the highest and lowest rankings.

| Chapters   | Top 3 Highest-Ranked Words |      | Bottom 3 Lowest-Ranked Words |      |
|------------|----------------------------|------|------------------------------|------|
|            | words                      | PoS  | words                        | PoS  |
| Chapter 1  | managing                   | vt.  | ascertain                    | vt.  |
|            | funding                    | vt.  | iterative                    | adj. |
|            | unify                      | vt.  | reimbursed                   | vt.  |
| Chapter 2  | sketch                     | n.   | disturbance                  | n.   |
|            | clay                       | n.   | conformance                  | n.   |
|            | precise                    | adj. | inherently                   | adv. |
| Chapter 3  | retrieve                   | vt.  | proprietary                  | n.   |
|            | outlay                     | n.   | subtracting                  | vt.  |
|            | classification             | n.   | wheelbarrows                 | n.   |
| Chapter 4  | tenet                      | n.   | unprecedented                | vt.  |
|            | codes                      | n.   | deliverables                 | n.   |
|            | audits                     | n.   | contamination                | n.   |
| Chapter 5  | tender                     | vi.  | pipelines                    | n.   |
|            | limits                     | n.   | subcontractors               | n.   |
|            | float                      | n.   | fluctuations                 | n.   |
| Chapter 6  | reward                     | n.   | investigatory                | adj. |
|            | insurance                  | n.   | jurisdictions                | n.   |
|            | programme                  | n.   | accusatorial                 | adj. |
| Chapter 7  | tracking                   | vt.  | unbiased                     | adj. |
|            | implementing               | vt.  | approximate                  | adj. |
|            | evaluating                 | vt.  | sagacious                    | adj. |
| Chapter 8  | rough                      | adj. | alternatives                 | n.   |
|            | thrust                     | n.   | externalities                | n.   |
|            | conduct                    | vt.  | interpretations              | n.   |
| Chapter 9  | provision                  | n.   | prosecutions                 | n.   |
|            | occurrence                 | n.   | consciousness                | n.   |
|            | bargaining                 | vi.  | inextricably                 | adv. |
| Chapter 10 | leader                     | n.   | drudgery                     | n.   |
|            | evoke                      | vt.  | metaphors                    | n.   |
|            | facets                     | n.   | conviction                   | n.   |
| Chapter 11 | operate                    | vt.  | taxonomies                   | n.   |
|            | drones                     | n.   | disambiguate                 | vt.  |
|            | ductwork                   | n.   | euphemistically              | adv. |
| Chapter 12 | developer                  | n.   | obsolescence                 | n.   |
|            | efficiency                 | n.   | shareholder                  | n.   |
|            | texture                    | n.   | multidisciplinary            | adj. |

**Table 8.** Phrase items with the highest and lowest rankings.

| Chapters   | Top 3 Highest-Ranked Phrases | Bottom 3 Lowest-Ranked Phrases |
|------------|------------------------------|--------------------------------|
|            | phrases                      | phrases                        |
| Chapter 1  | acts as                      | individual contractors         |
|            | approve plan                 | memorandum of understanding    |
|            | dry wall                     | schematic diagrams             |
| Chapter 2  | with the aid of              | naturally occurring substances |
|            | laying out                   | subsequent recipient           |
|            | comprising of                | drafting board                 |
| Chapter 3  | making up                    | concrete plants                |
|            | adding up                    | extension cords                |
|            | job site                     | to-date variances              |
| Chapter 4  | implement plans              | closure work                   |
|            | intrinsic merit              | disclose sth to sb             |
|            | holistic approach            | census taker                   |
| Chapter 5  | day-to-day                   | non-critical activities        |
|            | event date                   | suffice it to say              |
|            | time rage                    | de-conflicting                 |
| Chapter 6  | engaged in                   | periodic payments              |
|            | contract law                 | turnkey operation              |
|            | legal sanctions              | unbalanced bid                 |
| Chapter 7  | lose utility                 | planning risk responses        |
|            | risk management              | suspend execution              |
|            | tempt out                    | tempt strongly                 |
| Chapter 8  | makes use of                 | compounding operation          |
|            | by the time                  | simultaneous interpretation    |
|            | in principle                 | native asphalt                 |
| Chapter 9  | construction managers        | key performance indicators     |
|            | steel fixing                 | OHS competencies               |
|            | plan layouts                 | safety atmosphere surveys      |
| Chapter 10 | project leader               | interpersonal influence        |
|            | job function                 | constitute obstacles           |
|            | project participant          | trait identification           |
| Chapter 11 | the bulk of                  | hyper-competitive era          |
|            | delve into                   | be reverted for                |
|            | better customer service      | ameliorated connection         |
| Chapter 12 | total cost                   | landscape architect            |
|            | relevant costs               | from cradle to grave           |
|            | time frame                   | non-domestic buildings         |

Extending the semantico-syntactic dichotomy from discrete words to multi-word expressions, **Table 8** reveals a parallel cognitive mechanism governing phrase-level evaluations across the twelve chapters. The apex of the importance hierarchy is consistently dominated by phrases exhibiting high procedural tangibility and task-oriented salience (e.g., approve plan, risk management, total cost), which are deeply anchored in the procedural, managerial, and interactional domains of frontline engineering practice. Because the phrases carry immediate experiential salience, students can effortlessly map them onto observable professional workflows and daily site operations. Conversely, the lowest-ranked echelons are heavily populated by expressions characterized by institutional abstraction, idiomatic opacity, or heavy discourse dependence (e.g., memorandum of understanding, unbalanced bid, from cradle to grave) that operate predominantly within highly formalized, documentary, or conceptual registers. Rather than denoting immediate physical actions, they require broader discourse competence, field-specific conceptual inference, or familiarity with macro-level institutional frameworks. Upon the discussions above, the structural contrast corroborates that the pragmatic utility principle identified at the word level scales up to phrases due to students systematically prioritizing multi-word expressions with direct, operational deployability, while assigning lower importance to conceptually dense or idiomatic phrases that are decoupled from immediate, observable engineering realities.

Synthesizing the empirical evidence from both **Table 7** and **Table 8** evaluations, the analytical results emphasize that the IWAM offers profound methodological and pedagogical contributions across dual linguistic dimensions. By employing a systematic top 3 highest-ranked items and bottom 3 lowest-ranked items clustering analysis, the framework decodes the non-arbitrary, cognitive logic governing student judgments. It proves that student-perceived importance transcends mere statistical frequency, but also systematically mediates by operational tangibility, semantic transparency, and immediate vocational relevance across both individual words and complex expressions, while reshaping the pedagogical landscape of professional English. The IWAM in this study equips instructors with a scientifically grounded diagnostic tool to implement targeted, differentiated instruction, which allows both teachers and students to rapidly reinforce high-utility, task-oriented language while strategically scaffolding conceptually dense or discourse-heavy items. In conclusion, by seamlessly aligning instructional focus with the natural contours of learner cognition, the IWAM framework not only optimizes teaching efficiency but also significantly accelerates the overall trajectory of student acquisition, cementing its status as a highly actionable, process-oriented assessment paradigm.

## 5. Conclusion

This study conceptualized and validated the IWAM to decode the learning dynamics of professional English for engineering management, which captures classroom learning as a highly dynamic, non-linear process. By integrating students' perceived importance with pre- and post-class mastery shifts, IWAM trans-

lates subjective judgments into quantitative indicators. The analytical results note that the metric revealed a consistent semantico-syntactic dichotomy, demonstrating that students prioritize concrete, action-oriented items over dense or discourse-heavy phrases. Furthermore, IWAM exposed heterogeneous absorption trajectories, confirming that while discrete words exhibit structural stability, multi-word phrases demonstrate higher contextual volatility and a stronger reliance on specific instructional methods.

Also, this study substantiates the critical necessity for a process-oriented, student-centric evaluation paradigm within a pedagogical standpoint. Historically, language education has relied on reductionist summative examinations that obscure cognitive engagement and the iterative nature of knowledge assimilation (Abella et al., 2005; Erdem Coşgun, 2025; Fan & Jin, 2013; Foldnes, 2016; Ghorbandordinejad & Ahmadabad, 2016; Gu, 2014; Hatipoğlu, 2016; Lee & Wallace, 2018; Lu et al., 2025; Prediger et al., 2018; Sadeghi & Rahmati, 2017; Saritepeci et al., 2019; Shin, 2005; Tuomainen, 2018; Wharton, 2000; Woodrow, 2006; Yilmaz et al., 2022; Zheng & Bender, 2019). At the same time, IWAM provides a solution by embedding student-rated importance as a core evaluative layer, providing educators with a data-driven diagnostic tool for instructional triage, shifting the evaluation landscape toward a highly responsive system that faithfully reflects how students actualize and absorb professional English.

Despite its methodological rigor, this study acknowledges several structural constraints. The empirical validation was confined to a single 32-hour teaching cycle involving a purposive cohort of 80 undergraduate engineering management students, limiting broader extrapolation. Moreover, the observed results must be interpreted cautiously because this study utilized a one-group pretest-posttest design without a parallel control group. Due to this, the upward shifts in importance perception and linguistic mastery represent a verification of within-cohort growth and longitudinal trajectory alignment, rather than directional causal proof of external instructional superiority over alternative methodologies.

To address these limitations, future studies are recommended to employ fully randomized controlled trials (RCTs) or quasi-experimental comparative designs to establish explicit causality and isolate unique instructional variance from confounding longitudinal factors. Moreover, expanding the participant baseline across diverse disciplines and multiple academic years will further fortify the metric's theoretical foundation and generalizability. In conclusion, the IWAM framework offers a pathway for student-perception-based evaluation systems, providing a process-oriented instrument that aligns assessment methodologies with students' authentic cognitive trajectories.

## **Acknowledgements**

This work was supported by the Project of the Shaanxi Provincial Educational Science Planning Program, "Research on Dynamic Modeling of Undergraduate Students' Learning Behaviors and Academic Performance Prediction Based on

Artificial Intelligence”, under Grant AK-25-30, and the 2025 science and technology project of Ankang Municipal Science and Technology Bureau, “AI-based Plush Toy Image Generation and Design Using Deep Learning Neural Networks”, under Grant AK2025-GY-24.

Additionally, this work benefited significantly from the comprehensive resources provided by the Ankang Science & Technology Innovation Center (Qinchuangyuan Platform), which facilitated the execution and provided technical assistance for this scientific investigation.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- Abella, R., Urrutia, J., & Shneyderman, A. (2005). An Examination of the Validity of English-Language Achievement Test Scores in an English Language Learner Population. *Bilingual Research Journal*, 29, 127-144. <https://doi.org/10.1080/15235882.2005.10162827>
- Arana, S., Hagoort, P., Schoffelen, J. M., & Rabovsky, M. (2024). Perceived Similarity as a Window into Representations of Integrated Sentence Meaning. *Behavior Research Methods*, 56, 2675-2691. <https://doi.org/10.3758/s13428-023-02129-x>
- Beavers, J., Everdell, M., Jerro, K., Kauhanen, H., Koontz-Garboden, A., LeBovidge, E. et al. (2017). Two Types of States: A Cross-Linguistic Study of Change-of-State Verb Roots. *Proceedings of the Linguistic Society of America*, 2, Article 38. <https://doi.org/10.3765/plsa.v2i0.4094>
- Bernaisch, T., Gries, S. T., & Heller, B. (2022). Theoretical Models and Statistical Modelling of Linguistic Epicentres. *World Englishes*, 41, 333-346. <https://doi.org/10.1111/weng.12580>
- Bireta, T. J., & Mazzei, C. M. (2016). Does the Isolation Effect Require Attention? *Memory & Cognition*, 44, 1-14. <https://doi.org/10.3758/s13421-015-0538-y>
- Bruijnzeel, H., Cattani, G., Stegeman, I., Topsakal, V., & Grolman, W. (2017). Incorporating Ceiling Effects during Analysis of Speech Perception Data from a Paediatric Cochlear Implant Cohort. *International Journal of Audiology*, 56, 550-558. <https://doi.org/10.1080/14992027.2017.1311029>
- Bygate, M. (2016). Sources, Developments and Directions of Task-Based Language Teaching. *The Language Learning Journal*, 44, 381-400. <https://doi.org/10.1080/09571736.2015.1039566>
- Chen, X., Li, J., & Zhu, S. (2021). Translanguaging Multimodal Pedagogy in French Pronunciation Instruction: Vis-à-Vis Students’ Spontaneous Translanguaging. *System*, 101, Article 102603. <https://doi.org/10.1016/j.system.2021.102603>
- Cowan, N. (2014). Working Memory Underpins Cognitive Development, Learning, and Education. *Educational Psychology Review*, 26, 197-223. <https://doi.org/10.1007/s10648-013-9246-y>
- de Graaff, S., Hasselman, F., Bosman, A. M. T., & Verhoeven, L. (2008). Cognitive and Linguistic Constraints on Phoneme Isolation in Dutch Kindergartners. *Learning and Instruction*, 18, 391-403. <https://doi.org/10.1016/j.learninstruc.2007.08.001>
- Dudău, D. P., & Sava, F. A. (2021). Performing Multilingual Analysis with Linguistic Inquiry and Word Count 2015 (LIWC2015). An Equivalence Study of Four Languages.

- Frontiers in Psychology*, 12, Article ID: 570568.  
<https://doi.org/10.3389/fpsyg.2021.570568>
- Erdem Coşgun, G. (2025). Artificial Intelligence Literacy in Assessment: Empowering Pre-service Teachers to Design Effective Exam Questions for Language Learning. *British Educational Research Journal*, 51, 2340-2357. <https://doi.org/10.1002/berj.4177>
- Evans, N. W., Hartshorn, K. J., Cox, T. L., & De Jel, T. M. (2014). Measuring Written Linguistic Accuracy with Weighted Clause Ratios: A Question of Validity. *Journal of Second Language Writing*, 24, 33-50. <https://doi.org/10.1016/j.jslw.2014.02.005>
- Fan, J., & Jin, Y. (2013). A Survey of English Language Testing Practice in China: The Case of Six Examination Boards. *Language Testing in Asia*, 3, Article No. 7. <https://doi.org/10.1186/2229-0443-3-7>
- Foldnes, N. (2016). The Flipped Classroom and Cooperative Learning: Evidence from a Randomised Experiment. *Active Learning in Higher Education*, 17, 39-49. <https://doi.org/10.1177/1469787415616726>
- Gabaldón-Estevan, D. (2020). Heterogeneity versus Homogeneity in Schools: A Study of the Educational Value of Classroom Interaction. *Education Sciences*, 10, Article 335. <https://doi.org/10.3390/educsci10110335>
- Ghorbandordinejad, F., & Ahmadabad, R. M. (2016). Examination of the Relationship between Autonomy and English Achievement as Mediated by Foreign Language Classroom Anxiety. *Journal of Psycholinguistic Research*, 45, 739-752. <https://doi.org/10.1007/s10936-015-9371-5>
- Gu, L. (2014). At the Interface between Language Testing and Second Language Acquisition: Language Ability and Context of Learning. *Language Testing*, 31, 111-133. <https://doi.org/10.1177/0265532212469177>
- Hammers, D. B., Bothra, S., Polsinelli, A., Apostolova, L. G., & Duff, K. (2024). Evaluating Practice Effects across Learning Trials—Ceiling Effects or Something More? *Journal of Clinical and Experimental Neuropsychology*, 46, 630-643. <https://doi.org/10.1080/13803395.2024.2400107>
- Hatipoğlu, Ç. (2016). The Impact of the University Entrance Exam on EFL Education in Turkey: Pre-Service English Language Teachers Perspective. *Procedia-Social and Behavioral Sciences*, 232, 136-144. <https://doi.org/10.1016/j.sbspro.2016.10.038>
- Helal, S., Li, J., Liu, L., Ebrahimie, E., Dawson, S., Murray, D. J. et al. (2018). Predicting Academic Performance by Considering Student Heterogeneity. *Knowledge-Based Systems*, 161, 134-146. <https://doi.org/10.1016/j.knosys.2018.07.042>
- Hosseini, H., Hesar, F., & Marvasti, F. (2014). Real-Time Impulse Noise Suppression from Images Using an Efficient Weighted-Average Filtering. *IEEE Signal Processing Letters*, 22, 1050-1054. <https://doi.org/10.1109/lsp.2014.2381649>
- Hyland, K. (2019). English for Specific Purposes: Some Influences and Impacts. In X. Gao (Ed.), *Second Handbook of English Language Teaching* (pp. 337-353). Springer International Publishing. [https://doi.org/10.1007/978-3-030-02899-2\\_19](https://doi.org/10.1007/978-3-030-02899-2_19)
- Hyönä, J., Vainio, S., & Laine, M. (2002). A Morphological Effect Obtains for Isolated Words but Not for Words in Sentence Context. *European Journal of Cognitive Psychology*, 14, 417-433. <https://doi.org/10.1080/09541440143000131>
- Jassem, Z. A. (2012). The Arabic Origins of Common Religious Terms in English: A Lexical Root Theory Approach. *International Journal of Applied Linguistics & English Literature*, 1, 59-71. <https://doi.org/10.7575/ijalel.v.1n.6p.59>
- Jiang, J., Bi, P., & Liu, H. (2019). Syntactic Complexity Development in the Writings of EFL Learners: Insights from a Dependency Syntactically-Annotated Corpus. *Journal of Sec-*

- ond Language Writing*, 46, Article 100666. <https://doi.org/10.1016/j.jslw.2019.100666>
- Joshi, H., Rosenbloom, P. S., & Ustun, V. (2014). Isolated Word Recognition in the Sigma Cognitive Architecture. *Biologically Inspired Cognitive Architectures*, 10, 1-9. <https://doi.org/10.1016/j.bica.2014.11.001>
- Lee, G., & Wallace, A. (2018). Flipped Learning in the English as a Foreign Language Classroom: Outcomes and Perceptions. *TESOL Quarterly*, 52, 62-84. <https://doi.org/10.1002/tesq.372>
- Lu, J., Ma, Q., & Li, S. (2025). Effect of Localized Task-Based Language Teaching on Chinese Secondary School English Learners' Oral Production in Examination-oriented Contexts. *International Journal of Applied Linguistics*, 35, 168-192. <https://doi.org/10.1111/ijal.12608>
- Marcu, N. A. (2020). Designing Functional ESP (English for Specific Purposes) Courses. *Procedia Manufacturing*, 46, 308-312. <https://doi.org/10.1016/j.promfg.2020.03.045>
- Papadakis, N., Mémin, E., Cuzol, A., & Gengembre, N. (2010). Data Assimilation with the Weighted Ensemble Kalman Filter. *Tellus A: Dynamic Meteorology and Oceanography*, 62, 673-697. <https://doi.org/10.1111/j.1600-0870.2010.00461.x>
- Prediger, S., Wilhelm, N., Büchter, A., Gürsoy, E., & Benholz, C. (2018). Language Proficiency and Mathematics Achievement: Empirical Study of Language-Induced Obstacles in a High Stakes Test, the Central Exam ZP10. *Journal für Mathematik-Didaktik*, 39, 1-26. <https://doi.org/10.1007/s13138-018-0126-3>
- Rasmussen, L. S., Larsen, K., Houx, P., Skovgaard, L. T., Hanning, C. D., & Moller, J. T. (2001). The Assessment of Postoperative Cognitive Function. *Acta Anaesthesiologica Scandinavica*, 45, 275-289. <https://doi.org/10.1034/j.1399-6576.2001.045003275.x>
- Reid, R. D. (2015). *Theatre as TBLT: The Implementation of Theatre in a High School EFL Oral Communication Course in Japan*. Doctoral Dissertation, Victoria University of Wellington.
- Robinson, C. W., Best, C. A., Deng, W. (., & Sloutsky, V. M. (2012). The Role of Words in Cognitive Tasks: What, When, and How? *Frontiers in Psychology*, 3, Article ID: 95. <https://doi.org/10.3389/fpsyg.2012.00095>
- Robinson, P. (2011). Task-Based Language Learning: A Review of Issues. *Language Learning*, 61, 1-36. <https://doi.org/10.1111/j.1467-9922.2011.00641.x>
- Saadatmand, M., & Kumpulainen, K. (2012). Emerging Technologies and New Learning Ecologies: Learner's Perceptions of Learning in Open and Networked Environments. *Proceedings of the International Conference on Networked Learning*, 8, 266-275. <https://doi.org/10.54337/nlc.v8.9097>
- Sadeghi, K., & Rahmati, T. (2017). Integrating Assessment as, for, and of Learning in a Large-Scale Exam Preparation Course. *Assessing Writing*, 34, 50-61. <https://doi.org/10.1016/j.asw.2017.09.003>
- Sáez, J. A., Villacorta, P., & Corchado, E. (2019). Dataset Weighting via Intrinsic Data Characteristics for Pairwise Statistical Comparisons in Classification. In H. Pérez García, L. Sánchez González, M. Castejón Limas, H. Quintián Pardo, & E. Corchado Rodríguez (Eds.), *International Conference on Hybrid Artificial Intelligence Systems* (pp. 61-72). Springer International Publishing. [https://doi.org/10.1007/978-3-030-29859-3\\_6](https://doi.org/10.1007/978-3-030-29859-3_6)
- Saritepeci, M., Duran, A., & Ermiş, U. F. (2019). A New Trend in Preparing for Foreign Language Exam (YDS) in Turkey: Case of Whatsapp in Mobile Learning. *Education and Information Technologies*, 24, 2677-2699. <https://doi.org/10.1007/s10639-019-09893-4>
- Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal,

- M. et al. (2013). Personality, Gender, and Age in the Language of Social Media: The Open-Vocabulary Approach. *PLOS ONE*, 8, e73791. <https://doi.org/10.1371/journal.pone.0073791>
- Shadrova, A. (2025). No Three Productions Alike: Lexical Variability, Situated Dynamics, and Path Dependence in Task-Based Corpora. *Open Linguistics*, 11, Article 20240036. <https://doi.org/10.1515/opli-2024-0036>
- Shin, S. K. (2005). Did They Take the Same Test? Examinee Language Proficiency and the Structure of Language Tests. *Language Testing*, 22, 31-57. <https://doi.org/10.1191/0265532205lt296oa>
- Tan, Q., Dong, X., Li, Q., & Ren, Z. (2018). Distributed Event-Triggered Cubature Information Filtering Based on Weighted Average Consensus. *IET Control Theory & Applications*, 12, 78-86. <https://doi.org/10.1049/iet-cta.2017.0575>
- Tuomainen, S. (2018). Examination as the Method in the Recognition of Prior Language Learning. *International Journal of Lifelong Education*, 37, 676-688. <https://doi.org/10.1080/02601370.2018.1518346>
- Turmezei, T. D. (2012). The Linguistic Roots of Modern English Anatomical Terminology. *Clinical Anatomy*, 25, 1015-1022. <https://doi.org/10.1002/ca.22062>
- Tyng, C. M., Amin, H. U., Saad, M. N. M., & Malik, A. S. (2017). The Influences of Emotion on Learning and Memory. *Frontiers in Psychology*, 8, Article ID: 235933. <https://doi.org/10.3389/fpsyg.2017.01454>
- Vinje, H., Brovold, H., Almøy, T., Frøslie, K. F., & Sæbø, S. (2021). Adapting Statistics Education to a Cognitively Heterogeneous Student Population. *Journal of Statistics and Data Science Education*, 29, 183-191. <https://doi.org/10.1080/26939169.2021.1928573>
- Wharton, G. (2000). Language Learning Strategy Use of Bilingual Foreign Language Learners in Singapore. *Language Learning*, 50, 203-243. <https://doi.org/10.1111/0023-8333.00117>
- Woodrow, L. (2006). Anxiety and Speaking English as a Second Language. *RELC Journal*, 37, 308-328. <https://doi.org/10.1177/0033688206071315>
- Xu, Y., & Yu, Y. (2025). Representation and Processing of L2 Compositional Multiword Sequences: Effects of Token Frequency, Type Frequency, and Constituency. *Behavioral Sciences*, 15, Article 734. <https://doi.org/10.3390/bs15060734>
- Yilmaz, R. M., Topu, F. B., & Takkaç Tular, A. (2022). An Examination of the Studies on Foreign Language Teaching in Pre-School Education: A Bibliometric Mapping Analysis. *Computer Assisted Language Learning*, 35, 270-293. <https://doi.org/10.1080/09588221.2019.1681465>
- Yitzhak, N., Harel, A., Yaari, M., Friedlander, E., & Yirmiya, N. (2016). The Mullen Scales of Early Learning: Ceiling Effects among Preschool Children. *European Journal of Developmental Psychology*, 13, 138-151. <https://doi.org/10.1080/17405629.2015.1073584>
- Zaninello, A., & Birch, A. (2020). Multiword Expression Aware Neural Machine Translation. In N. Calzolari, F. Béchet, P. Blache et al. (Eds.), *Proceedings of the Twelfth Language Resources and Evaluation Conference* (pp. 3816-3825). European Language Resources Association. <https://aclanthology.org/2020.lrec-1.471/>
- Zhao, W. (2020). Epistemological Flashpoint in China's Classroom Reform: (how) Can a "Confucian Do-after-Me Pedagogy" Cultivate Critical Thinking? *Journal of Curriculum Studies*, 52, 101-117. <https://doi.org/10.1080/00220272.2019.1641844>
- Zheng, M., & Bender, D. (2019). Evaluating Outcomes of Computer-Based Classroom Testing: Student Acceptance and Impact on Learning and Exam Performance. *Medical Teacher*, 41, 75-82. <https://doi.org/10.1080/0142159x.2018.1441984>