

Needs for English for Specific Purposes Course Development: English Language Requirements in Global Textile Supply Chain Management

Xiaomei Li, Baik Zhang

Foreign Languages School, Luoyang Institute of Science and Technology, Luoyang 471023, Henan, China

Corresponding author: Baik Zhang (e-mail: ly1glw@163.com)

ABSTRACT The English usage scenarios in global supply chain management are highly specialized and context-dependent, resulting in a mismatch between existing English course content and actual workplace language requirements. Based on corpus linguistics and task-based needs analysis theory, this study develops a demand analysis framework that integrates language data-driven and task-oriented principles to transform industry language requirements into curriculum objectives through systematic job-task identification. The results show that communication competence accounts for 37.9% of the language requirements in procurement positions, while comprehension competence reaches 46.8% in production positions, demonstrating significant differences in language priorities across supply chain functions. Furthermore, the analysis reveals distinctive syntactic structures and modal expression patterns in procurement, production, and sales discourse, illustrating how conditional constructions, obligation modals, and evaluative expressions support different communicative purposes in professional contexts. By providing an objective and structured linguistic validation mechanism, the proposed framework establishes an evidence-based pathway for ESP curriculum development and offers methodological support for semantic information organization, intelligent communication, and knowledge transmission in digitally interconnected engineering systems, including electromagnetic information networks where accurate context-aware information exchange is increasingly essential.

INDEX TERMS global supply chain management, corpus linguistics, task-based needs analysis, semantic information processing, electromagnetic information networks

I. INTRODUCTION

THE global textile supply chain is a complex, cross-regional network involving procurement, production, transportation, and sales. Standardized English communication is essential for collaboration and decision-making in international trade [1,2]. However, current university ESP courses focus primarily on general language skills and fail to address the specific industry contexts and occupational language needs of the global textile sector [3,4]. Addressing this gap is crucial for enhancing textile professionals' global communication competence and competitiveness.

The diversity of job tasks and the complexity of industrial corpora pose major challenges for research on English instruction in the global textile supply chain [5,6]. Language functions differ significantly across links, ranging from technical specifications in production to compliance documentation in transportation, customer communication in sales, and negotiation in procurement. Traditional needs analysis methods fail to capture these systematic linguis-

tic features, leading to fragmented course content, limited functional coverage, and weakly structured, unmeasurable teaching objectives [7,8].

To address the above difficulties, the academic community has attempted to apply methods such as corpus linguistics, register analysis, and task-based demand analysis into ESP course development [9,10]. Corpus linguistics reveals high-frequency industry vocabulary and collocation patterns through large-scale text mining [11]. Register analysis constructs industry language feature models based on grammatical variables [12]. Task-based demand analysis identifies language functions and behavioral structures based on professional tasks [13]. These methods have improved the objectivity and data level of demand analysis to a certain extent, but there are still problems such as scattered corpus sources, incomplete task mapping, and insufficient accuracy in language function identification. A systematic analysis mechanism from industry context to course design has not yet been formed [14,15].

To bridge the gap between industry language needs and course content, this study proposes a system model integrating corpus linguistics with task-based needs analysis to transform industrial language data into a teaching framework. Using the global textile supply chain as context, authentic corpora from procurement, production, transportation, and sales are collected, processed, and analyzed through keyword, collocation, and register analyses. A task-language correspondence matrix is then constructed based on job and process documentation, enabling the systematic integration of industrial language patterns into ESP course design.

II. MODEL CONSTRUCTION FOR ENGLISH DEMAND IN THE GLOBAL TEXTILE SUPPLY CHAIN

A. APPROACH AND PROCESS

This study aims to develop a systematic needs analysis model for curriculum design and industry application by integrating corpus linguistics with task-based needs analysis [16,17]. The research proceeds through corpus collection, language analysis, and functional identification within an industry context, ultimately producing a curriculum-oriented knowledge mapping model. This framework combines the applied value of task-based learning with the empirical strengths of corpus-based approaches, as illustrated in Figure 1.

By integrating a corpus analysis path with a task identification path, this approach provides bidirectional support for needs identification, linking industrial context input to curriculum framework development. The corpus path supplies measurable linguistic features and register patterns, while the task path ensures alignment between professional activities and language functions through task-language mapping. These paths converge at a language function matrix, producing a curriculum demand matrix that generates the ESP framework and forms a closed loop from corpus analysis to tasks, functions, and curriculum design.

B. INDUSTRY LANGUAGE DATA COLLECTION AND PROCESS

This study collects industry language data by screening public corpora and integrating them with self-constructed corpora to ensure data authenticity, legitimacy, and industry representativeness. The corpus covers procurement, production, transportation, and sales. The information in this course originated from the World Trade Organization, the International Organization for Standardization, the United Nations Conference on Trade and Development, the International Textile Manufacturers Federation, the American Apparel & Footwear Association, the European Union Legal database, and from documents such as industry reports, standards documents, manuals for suppliers, trade contracts, logistics documents, and company operation manuals published by Maersk.

To ensure regional representativeness, the corpus incorporates enterprise documents and trade communications from

major textile centers in Asia, Europe, and North America. Data sources include corporate correspondence and training manuals from China and India, supplier communications and audit documents from Germany, Italy, and Turkey, and procurement and shipment records from the United States and Canada. All data were linguistically verified and region-tagged to preserve native register features and support cross-regional comparison, enhancing the corpus's representation of the global textile supply chain.

Besides the formal and institutional materials, the study includes a sampling of semi-authentic communication materials drawn from relationships with textile firms. These materials consist of snippets of correspondence, replies to inquiries, and short transactional messages that embody real patterns of discourse within the day-to-day context of working internally with textiles. As semi-authentic communication materials, their incorporation transmits descriptive and shared language data, diversifying the register coverage and representation of workplace communication. The texts in the corpus are all systematically formatted and normalized as a separate research corpus.

To ensure alignment with authentic communicative practices in the textile industry, a validation stage was added after corpus collection. Six industry professionals and two textile company employees reviewed corpus samples to assess whether discourse patterns, collocations, and communicative purposes reflected actual workplace use. Validated segments were retained as benchmark data and integrated into task mapping, ensuring the corpus represents authentic occupational communication rather than purely institutional documentation and thereby enhancing its practical validity.

All enterprise-derived textual materials used in this study were accessed through institutional-level cooperation or publicly available channels and were approved for academic research use prior to collection. Any communication records, internal documents, or training materials originating from enterprises were processed in an anonymized form before inclusion in the corpus, with all identifiers of individuals and organizations removed at the text preparation stage. The validation participants involved in corpus review contributed on a voluntary basis and were asked to evaluate linguistic patterns rather than personal or organizational content. No personal data, sensitive information, or private correspondence was collected or analyzed. The study followed institutional guidelines for social science research ethics and, due to the absence of human subject intervention or identifiable data, did not require formal ethical review approval. All corpus data were used solely for research purposes and are not redistributed beyond aggregated analytical results.

Corpus processing followed three stages: text extraction, data cleaning, and structural annotation. Source documents were converted to UTF-8 plain text with non-core elements removed [18,19]. Data cleaning included symbol normalization using regular expressions, lemmatization, and stopword removal to reduce semantic noise [20,21]. A two-stage



FIGURE 1. Flowchart of English demand analysis for the global textile supply chain

stop-word filtering strategy was applied, beginning with a standard English stop-word list from the Natural Language Toolkit to minimize the loss of domain-relevant terms. Each term was then inspected against the frequency distribution of the textile supply chain corpus and selectively added back to the stop-word list when its removal would reduce domain relevance, based on its relative frequency within the corpus. To maintain domain representativeness, key business terms (e.g., order, shipment, fabric, supplier) were retained during stop-word filtering. A validated stop list of 517 items was developed by two linguists using a 10,000-token sampling test, ensuring no business-critical terms were removed. The stop list and validation records were archived with corpus metadata for replicability. Structural annotation then organized the corpus by supply chain stages, with each document tagged for year, word count, category, and source [22,23]. The final corpus contains 1,986,000 words, distributed across procurement (26%), production (33%), transportation (21%), and sales (20%).

Keyword extraction adopts comparative corpus statistical method [24,25], and the formula is:

$$K_i = \log_2 \frac{f_i/N_t}{c_i/N_r}, \quad (1)$$

where f_i is the frequency of occurrence of the i -th term in the target corpus; c_i is the frequency of occurrence of the same term in the reference corpus; N_t is the total number of terms in the target corpus and N_r is the total number of terms in the reference corpus. Keyness is measured using log ratio as an indicator of effect size, and terms with a log ratio value above 3.84 are treated as salient keywords reflecting substantial distributional difference between the target and reference corpora. The reference corpus used is the business subset of the Corpus of Contemporary American English (COCA) to ensure register comparability and statistical validity. This subset contains nearly 12 million words from business news, trade journals, corporate documents, and

online communications published between 1990 and 2019, covering multiple business registers. Accessed via the official COCA interface in March 2025, keyword extraction was based on standardized frequency lists, ensuring reproducible keyness calculations and stable cross-register comparison.

Collocation analysis measures the aggregation strength between words through the mutual information coefficient [26,27], and the calculation formula is:

$$MI(x, y) = \log_2 \left(\frac{P(x, y)}{P(x)P(y)} \right), \quad (2)$$

$P(x, y)$ is the probability of co-occurrence of terms x and y within a five-word window, and $P(x)$ and $P(y)$ are the probabilities of independent occurrence, respectively. Word pairs with a mutual information value of 3 or higher are considered stable collocations, constituting the semantic core structure of industry expressions. Based on the statistical results, a collocation network is constructed to reveal the functional associations and expression patterns of language in various links of the supply chain.

Register feature analysis adopts Biber's multidimensional model, applying principal component analysis to fifteen grammatical variables [28,29]. Normalized Z-scores are used to map texts onto a two-dimensional information-density/interaction-level space, revealing functional variation across supply chain stages. Procurement and transportation texts favor factual, descriptive language, whereas production and sales exhibit more interactive and communicative patterns.

Keywords, collocations, and grammatical features are vectorized during the comprehensive modeling step to create a language feature matrix $L = [w_{ij}]$, where w_{ij} is the normalized weight of the j -th language feature in the i -th task context. Cosine similarity is used in semantic clustering to determine the functional closeness and aggregation links of language features. A quantitative foundation for mapping work activities to linguistic attributes is provided by the

clustering results, which create clusters of procurement negotiation expressions, production process instructions, transportation compliance records, and sales communication statements. In order to create quantifiable expression of industry language guidelines, this procedure is data-driven. The results of the study are objective and repeatable, providing compelling evidence for creating learning objectives and subsequently revising course requirements.

C. POSITION TASK IDENTIFICATION AND COURSE REQUIREMENTS CONVERSION MECHANISM

Position identification is based on task-based requirements analysis. It systematically transforms linguistic needs into learning objectives through the use of functional modeling, corpus mapping, and task decomposition. Production, sales, transportation, and procurement are the four employment domains that are examined in this study. Training manuals, performance standards, business procedure documentation, and job descriptions are a few examples of data sources. A task-language mapping system is built using the corpus analysis results.

Using a hierarchical decomposition technique, position identification breaks down job activities into three levels: operational stages, subtasks, and goal tasks [30,31]. A hierarchical encoding technique is used to include this structure into the corpus indexing system, allowing for an accurate mapping of job context and language resources.

To measure the concentration of language features in various task contexts, task language relevance is computed during the corpus mapping step using the language feature matrix [32,33]. The calculation formula is:

$$R_i = \frac{\sum_{j=1}^n w_{ij} \cdot s_j}{\sum_{j=1}^n s_j}, \quad (3)$$

R_i is the linguistic relevance of the job, and s_j is the frequency of occurrence of the linguistic feature throughout the entire corpus.

The weight parameter w_{ij} represents the normalized contribution of a given linguistic feature to a specific task context and is derived through corpus-automatic mapping between annotated task segments and language feature vectors. The initial assignment of w_{ij} values is computed based on relative feature concentration within task-aligned subcorpora, followed by expert validation of task-feature alignment. Inter-rater consistency was assessed during the validation stage, and the resulting agreement level ensured the stability of the weight assignment used in subsequent calculations.

A higher relevance score indicates a greater reliance on the relevant linguistic component for the job. Based on register theory, language functions are separated into four groups during the functional modeling stage: behavioral instructions, relationship maintenance, information transmission, and evaluative remarks. Four functional categories were defined using lexicosemantic and pragmatic criteria

based on Halliday's metafunctions. These include behavioral instruction, relationship maintenance, information transmission, and evaluative remark, distinguished by their dominant illocutionary force and linguistic markers. Sentence-level functional coding was conducted through semi-automatic tagging in UAM Corpus Tool and subsequently validated by an ESP register specialist. The final inter-coder reliability was assessed using Cohen's Kappa coefficient, which reached a value of 0.89 after manual validation, indicating a high level of agreement beyond chance and supporting the replicability of the functional classification results.

Through the statistical analysis of verb phrases, modal expressions, cohesive words and discourse structure patterns, task-function matrix is obtained to show the relationship between task behaviors and language functions. The results show that relationship maintenance and evaluative statements are dominant in sales discourse, behavioral instructions are most frequent in purchasing and production tasks, while information transmission constitutes the primary function in transportation-related work, indicating that positional attributes systematically constrain the distribution of language functions.

The course requirements derivation multiply the language feature matrix by the task-function matrix to yield a course requirements matrix on which the mix of language elements and the hierarchy of arrangement of the instructional modules [34,35] are based. The course system comprises four modules: procurement communication, production coordination, transportation records, and sales presentations, each of which is centered on a task language function. The language implementation level guides the teaching objectives, thus attaining a match between the level of job behavior and the level of course content.

The conversion function translates linguistic relevance values and task-based functional proportions into concrete curriculum parameters. Normalized relevance scores are aggregated by function to determine instructional weights, which are then used to allocate teaching hours and learning units across modules. Expert-calibrated constants balance linguistic significance with pedagogic feasibility, while functional proportions are mapped to learning outcomes and assessment criteria. This process establishes a measurable link between corpus analysis, instructional design, and evaluative focus, enabling the course requirement matrix to guide module sequencing, learning intensity, and assessment priorities.

All operational definitions of functional categories and representative annotated corpus excerpts used for validation are fully described within the methodological sections of this article to ensure transparency of the classification procedure and facilitate analytical replication. This is a mechanism based on corpus data and taking task functions as the main line. It converts industry language rules into operational teaching design elements, making the whole process of demand analysis have data rigidity and course orientation, which offers systematic support for the genera-

tion of course framework.

After the initial modelling, this paper held a panel discussion with managers and technical supervisors from the participating companies to review the first drafts of task-language mappings. This paper collected comments on whether the suspected communicative functions matched up with actual job behaviors and interactional routines, and revised descriptions of tasks and assignments of language functions when discrepancies were found. Thus, the validation process ensured that the final course requirement matrix reflected not only patterns found in the corpus, but also lived communicative demands in textile workplaces.

III. EMPIRICAL RESULTS ON ENGLISH LANGUAGE DEMAND IN THE GLOBAL TEXTILE SUPPLY CHAIN

A. LINGUISTIC CHARACTERISTICS OF INDUSTRY CORPUS DISTRIBUTION

This section is aimed at analysing the use of English in procurement, production, transportation and sales through the comparison of word frequencies in the industry corpus. The vocabulary function is studied and an eight-category classification system is developed based on the study: business instructions, relationship maintenance, process control, quality and standards, logistics and records, market and customers, evaluation and attitude, and management and coordination.

Each category was operationally defined using a corpus-driven coding framework refined through three rounds of expert consultation. The categories cover business instructions, relationship maintenance, process control, quality and standards, logistics and records, market and customers, evaluation and attitude, and management and coordination, each defined by its core lexical and functional features. All lexical tokens were automatically extracted and preliminarily tagged using keyword-frequency filtering, then manually verified by two trained coders. Inter-coder reliability was high, with a Cohen's kappa coefficient of 0.87, indicating robust classification consistency.

By calculating the distribution ratio of various vocabulary types in different links through corpus stratification, the aggregation trend and functional emphasis of industry language resources in the task context can be observed. The distribution characteristics of each link are shown in Figure 2:

As shown in Figure 2, business instruction and relationship maintenance vocabulary predominate in procurement discourse, with standardized frequencies of 26.8 and 22.6, indicating the transactional and collaborative nature of purchasing communication. Pragmatic analysis further shows that procurement discourse favors cooperative and reassuring expressions (e.g., confirm, appreciate), whereas sales discourse relies more on evaluative and persuasive language (e.g., recommend, advise), reflecting differing relationship-maintenance strategies. Manufacturing discourse exhibits higher frequencies of process control (31.7) and quality-related vocabulary (27.9), highlighting

its reliance on operational directives and technical specifications. Transportation shows the highest concentration of logistics and record-keeping vocabulary (39.6), underscoring documentation and compliance demands. In the sales phase, market and customer-related vocabulary (28.7) and relationship maintenance terms (24.1) dominate, indicating a customer-centered communicative focus. Overall, the distribution patterns demonstrate clear functional differentiation of language use across supply chain stages, supporting a task-oriented and semantically clustered structure of industry discourse. Comparative analysis of regional subcorpora shows stable communicative functions across major textile regions but contrasting linguistic realizations. Chinese and Indian data exhibit higher frequencies of procedural verbs and obligation modals, whereas European and North American corpora show greater use of evaluative and customer-oriented expressions, reflecting regional discourse norms. These findings confirm that global supply chain English maintains functional consistency while incorporating cross-regional linguistic variation, which is integrated into the proposed model.

B. RELATIONSHIP BETWEEN POSITION TASKS AND LANGUAGE FUNCTIONS

Continuing with the differentiated pragmatic realizations detected within the comparison of relationship maintenance occupational discourse, the subsequent functional analysis has expanded the comparison to grammatical and lexical items for occupational discourses. Continuing the analytical progression from lexical distributions to functional differentiation as illustrated in Figure 3, this section interprets the pragmatic tendencies identified in Figure 2 through their corresponding syntactic and modal realizations in position-specific discourse.

Based on corpus statistics and task mapping, this study analyzes English language requirements in the global textile supply chain by modeling the distribution of job-related language functions. Focusing on procurement, transportation, production, and sales, it examines four functional dimensions—information transmission, behavioral instruction, relationship maintenance, and evaluation—and constructs an occupational language function matrix. The results, illustrated in Figure 3, reveal systematic differences in language use across job categories and clarify the link between task characteristics and language skills.

Figure 3 shows that relationship maintenance and information transmission account for the majority of the language used in procurement jobs (33% and 31%, respectively). This distribution demonstrates how collaboration, dialogue, and information verification are essential to procurement. As demonstrated by the fact that 44% of the language used in production jobs is behavioral directions, and the language used in the production process is often directive and standardized, with an emphasis on operational execution and process control. The vocabulary employed in this process is widely established, with a focus on reporting,

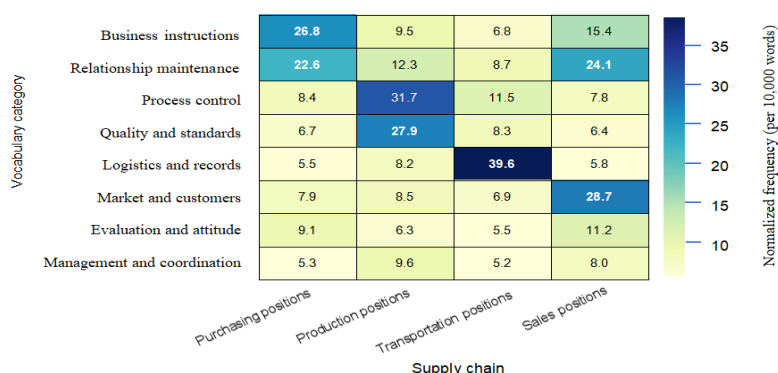


FIGURE 2. Heatmap of vocabulary category distribution in each segment

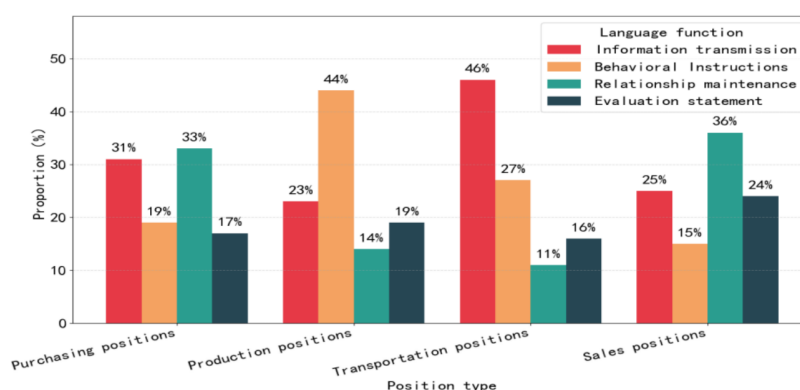


FIGURE 3. Position task-language function mapping

recording, and status descriptions, while information transmission (46%) is a component of transportation-related jobs. Relationship maintenance accounts for 36% and assessment remarks for 24%, reflecting the dual job emphasis of client contact and promotional statements in sales discourse. The whole structure provides empirical support for the function-oriented design of industrial English courses, showing that there is a stable relationship between language form and task qualities and significant difference in the language functions of different supply chain roles.

A further linguistic analysis identified the grammatical and lexical configurations underlying functional contrasts across positions. Procurement texts commonly employ conditional clauses and deontic modals. Production discourse features frequent passive constructions and non-finite verb groups expressing technical precision. Sales texts rely on epistemic modals and evaluative adjectives to project persuasion and assessment. These patterns indicate that grammatical differentiation across roles shapes communicative intent and professional tone.

C. CONSTRUCTION OF INDUSTRIAL ENGLISH NEEDS HIERARCHY MODEL

To avoid bias from preset corpus categories, an independent linguistic split was conducted before analyzing the three skill components. An unsupervised clustering algorithm grouped lexical and syntactic features according to communicative, expressive, or comprehension functions, independent of industrial roles. Only patterns converging across both unsupervised and task-based analyses were retained, ensuring that skill proportions were derived from language-internal evidence rather than predefined task divisions. The paper analyzes the quantity and proportion of different levels of language proficiency in procurement, transportation, production and sales jobs by using corpus data. Understanding, expression and communication are the three components of language proficiency. By calculating and scaling the data, a three-dimensional structural model of the three components is obtained, which reflects the differences in language requirements for jobs. The resulting hierarchical alignment of course content is based on the visualization of the weighted relationships between language proficiency levels. Figure 4 illustrates the relative composition of language proficiency across each link, reflecting the hierarchical differences in the language characteristics of industry tasks.

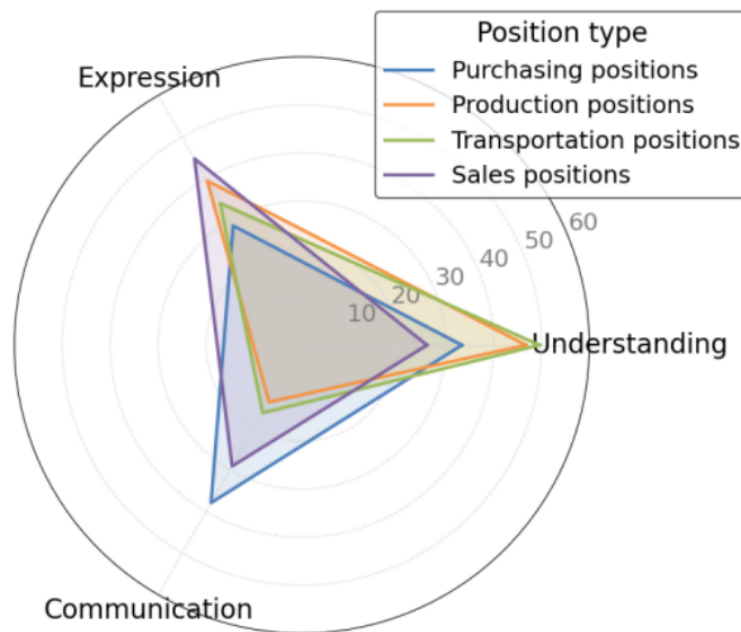


FIGURE 4. English language proficiency hierarchy distribution - global textile supply chain position comparison

As shown in Figure 4, the Purchasing position's communication skills account for 37.9%, significantly higher than the comprehension and expression skills (33.4% and 28.7%), indicating that language activities in the procurement process are centered around communication, negotiation, and relationship maintenance. The Production position's comprehension skills account for 46.8%, indicating a strong reliance on accurate interpretation of technical documents and process specifications within this role. Its expression skills (39.5%) are used for operational instructions and process feedback. The transportation position's comprehension skills account for 49.6%, significantly higher than its expression skills (34.1%) and communication skills (16.3%), reflecting the reliance of language tasks in this role on document reading and status recording. Communication skills make up 29.1% of the sales role, while expressive skills make up 44.8%. This implies that sales discourse focuses on customer communication and product promotion and has a more captivating and convincing linguistic structure. The tower distribution of language skills across professions confirms the analytical differentiation and functional complementarity of link, comprehension, express, and communicate skills. These distributions were tested against the results of the independent clustering test mentioned above, which showed that the respective skill emphases did not merely reproduce stratification by job type on the corpus but were reflective of statistically recurring linguistic tendencies among the full data set. This validation procedure enhances the analytical neutrality of the demand hierarchy

model and facilitates the interpretation of the results as evidence of objectively observed use of language rather than reiterations of prior task assumptions. This distribution typically closely fits the task attributes and the industry logic of the supply chain. Internal linguistic evidence supporting this distribution shows that the respective communicative emphases are recurring in terms of grammar and language rather than occupations alone, thus ensuring the functional hierarchy is grounded in objectively observed language form.

D. COURSE REQUIREMENTS FRAMEWORK GENERATION AND ALIGNMENT

This section assesses the course evaluation quantification of the alignment of course content to industry requirements in terms of four modules, namely, sales presentation, production coordination, shipping record, and procurement communication from the perspective of five evaluation indexes, namely, task coverage, corpus dependency, skill integration, instructional operability, and industry adaptation. By computing the matching coefficient based on standardized data of the four modules, several methods have been used to verify that the framework of the course is reasonable and balanced. Quantitative calculation of the five matching dimensions was carried out based on the standardized ratios shown in the course requirement matrix. This paper operationalized each dimension using a numerical indicator. Task coverage was calculated as the ratio of task elements in the course module that also appeared in the task-function matrix

extracted from the industry corpus. Skill integration was calculated as the weighted mean of the linguistic relevance scores of comprehension, expression, and communication skills, normalized to a 0–1 scale according to the functional weight of each skill in the task domain. Corpus dependency was calculated as the ratio of examples from the course originating from expressions validated by the corpus, to the total number of examples used for instruction in the module. Teaching operability was calculated as the ratio of task-language pairs converted into explicit classroom activities, to the total number of task-language pairs mapped from the course, normalized using a constant from the pilot course implementation feedback. Industry adaptability was calculated as the mean of expert ratings ($n = 8$) on the five-point scale of how closely they believed the content of the module reflected authentic practice in the industry with respect to the communication practices used, normalized to a 0–1 range. Finally, the matching value of each module was calculated as the mean of the five standardized indicators. Computation procedure 1, referring to the standardized calculation sequence defined through the normalization and aggregation of the five evaluation indicators described above in this section, provides a transparent and replicable basis for the values reported in Figure 5. The matching results for each module across several dimensions are shown in Figure 5, which also illustrates the demand-driven curriculum system's structural performance and distribution features.

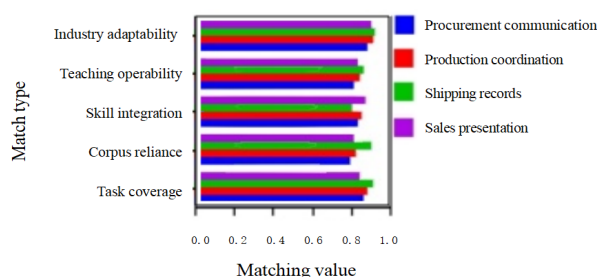


FIGURE 5. ESP course module matching results analysis

A high degree of alignment between the course design and industry needs is demonstrated by Figure 5, which shows that the overall fit of the four course modules stayed above 0.80. Its content is most closely matched with real-world business operations, and its corpus composition is most compatible with job activities, according to the Transportation Records module, which receives ratings of 0.91 and 0.92 in task coverage and industry adaptability, respectively. With a high of 0.91 and a fit that is uniformly spread throughout the five dimensions, the Production Coordination module demonstrates a strong alignment between operational requirements and expression skills. With a skill integration score of 0.87, the Sales Presentation module demonstrates that its instructional activities successfully combine professional communication with language expression. According to the module's fit, which varies from

0.79 to 0.88, the Procurement Communication module's course material consistently covers both communication and information processing. Overall, these findings show that the course framework developed in this study provides a quantitative foundation for the efficacy of ESP course development by achieving a high degree of congruence between industry task features and instructional design aspects.

IV. CONCLUSIONS

This study develops a model of English language needs in the global textile supply chain by integrating task-based analysis with corpus data. The results reveal clear functional differentiation across positions, with procurement emphasizing business instruction and relationship maintenance (26.8, 22.6), production relying on process control (31.7), and transportation dominated by logistics and record-keeping vocabulary (39.6). Job-specific distributions further confirm stable mappings between language functions and professional tasks, such as behavioral instruction in production (44%), information transmission in transportation (46%), and relationship maintenance in sales (36%). Evaluation of the curriculum needs framework demonstrates strong instructional validity, with all modules exceeding a matching degree of 0.80 and the transportation record-keeping module achieving 0.91–0.92 in task coverage and industry adaptability. Cross-verification between corpus analysis and practitioner validation confirms the authenticity of the identified linguistic patterns and supports task-based differentiation of ESP requirements at the occupational discourse level. Despite this alignment, the corpus remains limited in dynamic workplace interaction data; future research will expand enterprise-based data sources, incorporate semantic modeling and dynamic task contexts, and integrate regional linguistic variation to enhance the adaptability and practical relevance of ESP curriculum design in global textile supply chain contexts.

AUTHOR CONTRIBUTIONS

Conceptualization-Xiaomei Li and Baike Zhang; methodology-Xiaomei Li and Baike Zhang; formal analysis-Xiaomei Li and Baike Zhang; investigation-Xiaomei Li and Baike Zhang; resources-Xiaomei Li and Baike Zhang; writing-original draft preparation-Xiaomei Li and Baike Zhang; writing-review and editing-Xiaomei Li and Baike Zhang; supervision-Xiaomei Li and Baike Zhang. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

- [1] M. M. Roshid and A. Kankaanranta, "English communication skills in international business: Industry expectations versus university preparation," *Business and Professional Communication Quarterly*, vol. 88, no. 1, pp. 100–125, 2025, doi: 10.1177/23294906231184814.
- [2] F. Abbas, A. M. K. Rana, I. Bashir, and A. M. Bhatti, "The English language proficiency as a global employment skill: the viewpoint of Pakistani academia," *Humanities and Social Sciences Review*, vol. 9, no. 3, pp. 1071–1077, 2021, doi: 10.18510/hssr.2021.93106.
- [3] G. Shewangizaw and A. Hailu, "Investigation of English Language Needs of Trainers of Garment and Apparel Fashion Technology," *English Language, Literature & Culture*, vol. 8, no. 1, pp. 10–21, 2023, doi: 10.11648/j.ellc.20230801.13.
- [4] A. Munir, N. Darwis, L. D. D. Hidayat, and H. Nur, "A needs analysis of ESP for fashion design students at an Indonesian vocational school," *International Journal of Research on English Teaching and Applied Linguistics*, vol. 6, no. 1, pp. 24–37, 2025, doi: 10.30863/ijretal.v6i1.9787.
- [5] M. D. Islam and S. Ahmad, "English for the Apparel Merchandisers in Bangladesh: A Needs Analysis," *International Journal of Language Instruction*, vol. 4, no. 2, pp. 33–52, 2025, doi: 10.54855/ijli.25422.
- [6] G. Shewangizaw, A. Hailu, and W. Haile, "A comparative analysis of language features across communicative English skills and garment vocational texts," *Celtic: A Journal of Culture, English Language Teaching, Literature and Linguistics*, vol. 11, no. 2, pp. 353–376, 2024, doi: 10.22219/celtic.v11i2.35885.
- [7] F. Mao and J. Zhou, "A needs analysis of ESP courses in colleges of art and design: Consensus and divergence," *PLoS ONE*, vol. 19, no. 6, Art. no. e0305210, 2024, doi: 10.1371/journal.pone.0305210.
- [8] E. Petraki and K. Khat, "Challenges and constraints in the design of an ESP course in Cambodia: Implications for higher education institutions," *Asia Pacific Journal of Education*, vol. 42, no. 2, pp. 260–275, 2022, doi: 10.1080/02188791.2020.1798738.
- [9] N. Curry and T. McEnery, "Corpus linguistics for language teaching and learning: A research agenda," *Language Teaching*, vol. 58, no. 2, pp. 1–20, 2025, doi: 10.1017/S0261444824000430.
- [10] R. Syaripuddin, R. Nur, and A. Latifa, "Need analysis to develop English writing materials for ESP instruction," *Journal of English Education*, vol. 8, no. 1, pp. 23–30, 2023, doi: 10.31327/jee.v8i1.1901.
- [11] I. Xodabande, M. R. Atai, and M. R. Hashemi, "Analyzing large text data for vocabulary profiling in corpus-based studies of academic discourse," *MethodsX*, vol. 14, Art. no. 103168, 2025, doi: 10.1016/j.mex.2025.103168.
- [12] D. Biber, T. Larsson, and G. R. Hancock, "The linguistic organization of grammatical text complexity: Comparing the empirical adequacy of theory-based models," *Corpus Linguistics and Linguistic Theory*, vol. 20, no. 2, pp. 347–373, 2024, doi: 10.1515/cllt-2023-0016.
- [13] S. Keller and R. Gilabert, "Bienvenidos A Bordo: From Task-Based Needs Analysis to Design: Spanish-Destination Flight Attendants," *Canadian Journal of Applied Linguistics*, vol. 26, no. 2, pp. 81–108, 2023, doi: 10.37213/cjal.2023.32985.
- [14] A. Mudinillah, S. N. Rahmi, and N. Taro, "Task-based language teaching: A systematic review of research and applications," *Lingeduca: Journal of Language and Education Studies*, vol. 3, no. 2, pp. 102–115, 2024, doi: 10.70177/lingeduca.v3i2.1352.
- [15] P. A. P. Sudana, N. M. Ratminingsih, N. N. Padmadewi, and N. L. P. S. Adnyani, "Industry-Driven Task Needs and Syllabus Recommendation for Business English Materials," *International Journal of Language Education*, vol. 8, no. 4, pp. 734–748, 2024, doi: 10.26858/ijole.v8i4.70001.
- [16] T. T. Tran, T. T. Pham, and T. T. Le, "Integrating Corpus Linguistics in Vietnamese English Education: A Qualitative Inquiry into Scholarly Perspectives and Practices," *Higher Learning Research Communications*, vol. 14, no. 2, pp. 5, 2024, doi: 10.18870/hlrc.v14i2.1557.
- [17] A. Shadrova, "No three productions alike: Lexical variability, situated dynamics, and path dependence in task-based corpora," *Open Linguistics*, vol. 11, no. 1, Art. no. 20240036, 2025, doi: 10.1515/opli-2024-0036.
- [18] S. Avasthi, R. Chauhan, and D. P. Achariya, "Extracting information and inferences from a large text corpus," *International Journal of Information Technology*, vol. 15, no. 1, pp. 435–445, 2023, doi: 10.1007/s41870-022-01123-4.
- [19] M. H. A. Abdullah, N. Aziz, S. J. Abdulkadir, H. S. A. Alhussian, and N. Talpur, "Systematic literature review of information extraction from textual data: Recent methods, applications, trends, and challenges," *IEEE Access*, vol. 11, pp. 10535–10562, 2023, doi: 10.1109/ACCESS.2023.3240898.
- [20] C. P. Chai, "Comparison of text preprocessing methods," *Natural Language Engineering*, vol. 29, no. 3, pp. 509–553, 2023, doi: 10.1017/S1351324922000213.
- [21] A. P. Joshi and B. V. Patel, "Data preprocessing: The techniques for preparing clean and quality data for data analytics process," *Oriental Journal of Computer Science and Technology*, vol. 13, no. 2, pp. 78–81, 2021, doi: 10.13005/ojst13.0203.03.
- [22] P. Kalamkar, A. Tiwari, A. Agarwal, S. Karn, S. Gupta, V. Raghavan, et al., "Corpus for Automatic Structuring of Legal Documents," *arXiv e-prints*, Art. no. 2201.13125, 2022, doi: 10.48550/arXiv.2201.13125.
- [23] V. H. Baghdasaryan, "ArmSpeech: Armenian spoken language corpus," *International Journal of Scientific Advances*, vol. 3, no. 3, pp. 454–459, 2022, doi: 10.51542/ijscia.v3i3.25.
- [24] T. Nomoto, "Keyword extraction: A modern perspective," *SN Computer Science*, vol. 4, no. 1, pp. 92, 2022, doi: 10.1007/s42979-022-01481-7.
- [25] D. Sheng, "Statistics in corpus linguistics: A practical guide," *Social Semiotics*, vol. 33, no. 4, pp. 909–912, 2023, doi: 10.1080/10350330.2021.1969215.
- [26] Q. Su, C. Gu, and P. Liu, "Association measures for collocation extraction: Automatic evaluation on a large-scale corpus," *International Journal of Corpus Linguistics*, vol. 29, no. 1, pp. 59–86, 2024, doi: 10.1075/ijcl.21056.su.
- [27] M. Eguchi and K. Kyle, "L2 collocation profiles and their relationship with vocabulary proficiency: A learner corpus approach," *Journal of Second Language Writing*, vol. 60, Art. no. 100975, 2023, doi: 10.1016/j.jslw.2023.100975.
- [28] Y. Xu and Y. Zhang, "A Multi-Dimensional Analysis of Conclusions in Research Articles of Linguistics," *European Journal of Theoretical and Applied Sciences*, vol. 1, no. 6, pp. 191–203, 2023, doi: 10.59324/ejtas.2023.1(6).20.
- [29] S. Rashid, B. Mahmood, and Z. Sarfraz, "Register Variation in Pakistani English Political Discourse: An Instigation of Twitter Analytics Using Biber's Multidimensional Analysis," *International Journal of Academic Research for Humanities*, vol. 4, no. 2, pp. 210–230, 2024, [Online]. Available: <https://zenodo.org/records/12611875>.
- [30] L. Wang, J. Xie, X. Zhang, H. Su, and J. Zhu, "HIDE-PET: Continual learning via hierarchical decomposition of parameter-efficient tuning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 47, no. 8, pp. 6687–6702, 2025, doi: 10.1109/TPAMI.2025.3562534.
- [31] H. Hihn and D. A. Braun, "Hierarchically structured task-agnostic continual learning," *Machine Learning*, vol. 112, no. 2, pp. 655–686, 2023, doi: 10.1007/s10994-022-06283-9.
- [32] M. Deng, L. Tao, and J. Benton, "Measuring Feature Sparsity in Language Models," *arXiv e-prints*, Art. no. 2310.07837, 2023, doi: 10.48550/arXiv.2310.07837.
- [33] W. Huang, Y. Zhang, X. Zheng, F. Chao, and R. Ji, "Determining Layer-wise Sparsity for Large Language Models Through a Theoretical Perspective," *arXiv e-prints*, Art. no. 2502.14770, 2025, doi: 10.48550/arXiv.2502.14770.
- [34] D. Frost and R. Ackrill, "The multiple dimensions of curriculum mapping: Designing a comprehensive outcomes-based framework," *London Review of Education*, vol. 23, no. 1, pp. 17, 2025, doi: 10.14324/LRE.23.1.17.
- [35] B. Cooper, B. Cowie, and J. Furness, "Curriculum mapping as a boundary encounter: Meeting the demands of multiple agendas," *Educational Research for Policy and Practice*, vol. 23, no. 2, pp. 229–250, 2024, doi: 10.1007/s10671-021-09299-5.