

From Concordance to Confidence: A Sequential Explanatory Evaluation of a Digital Inductive Model for Linguistic Thinking in Pre-Service French Teacher Education

Mohamed Mekheimer

mohamed.abdelgawad3@edu.bsu.edu.eg

Beni-Suef University

Fatma Abdelaal

Beni-Suef University

Research Article

Keywords: data-driven learning, digital inductive model, French as a Foreign Language, linguistic thinking, pre-service teacher education, mixed methods, corpus pedagogy

Posted Date: August 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10093338/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

This sequential explanatory mixed-methods study examined whether a structured digital inductive model could strengthen linguistic thinking and reshape pedagogical orientations among pre-service French as a Foreign Language (FFL) teachers in Egypt. The model combined guided concordance work, a purpose-built corpus of authentic French texts, collaborative pattern discovery, and reflective discussion across a ten-week course intervention. Fifty-six third-year student-teachers completed the intervention; matched quantitative data were available for 55 participants. Linguistic thinking was assessed through a researcher-developed Linguistic Analysis Task (LAT) targeting pattern recognition, evidence-based hypothesis formulation, and metalinguistic reflection. Paired-samples analysis indicated a substantial observed increase in LAT performance from pre-test to post-test, $t(54) = 21.43$, $p < .001$, $d = 2.89$, with gains across all three subcomponents. Reflective journals and a focus group were then analyzed thematically to explain the quantitative pattern. Qualitative findings indicated a staged movement from initial uncertainty and technical friction toward active inquiry, confidence in using authentic evidence, and a more facilitative conception of language teaching. The results suggest that corpus-supported inductive work can offer a productive route for developing analytic language-teaching competence. Because the quantitative phase used a single-group pre-test/post-test design, the findings should be interpreted as strong evidence of within-cohort development rather than as definitive causal proof. The study contributes a refined model for linking data-driven learning, digital pedagogical competence, and pre-service teacher identity formation in technology-enhanced language teacher education.

Introduction

Language teacher education is increasingly expected to prepare future teachers for digitally mediated, evidence-rich, and analytically demanding classrooms. Yet the presence of digital tools in teacher preparation does not in itself guarantee pedagogically meaningful learning. A central challenge is therefore not whether pre-service teachers can operate technology, but whether they can use digital resources to investigate language, interpret authentic evidence, and design instruction that moves learners beyond rule memorization. This challenge is particularly relevant to grammar and lexis instruction, where deductive presentation remains common and where learners often experience language knowledge as fixed rules rather than as patterns of use that can be investigated.

The present study addresses this challenge through a digital inductive model for pre-service teachers of French as a Foreign Language (FFL). The model brings together corpus-based data-driven learning (DDL), guided inductive pedagogy, and collaborative reflection. Rather than receiving a grammatical rule and applying it to decontextualized examples, participants search a corpus of authentic French texts, compare concordance lines, identify recurrent patterns and exceptions, formulate provisional generalizations, and revise those generalizations through peer and instructor discussion. The pedagogical assumption is that such inquiry can develop linguistic thinking: the ability to notice patterns, reason from language evidence, formulate hypotheses, and explain form-meaning relationships in context.

This construct is important because language teachers need more than declarative knowledge of grammar. They must make analytical decisions while planning lessons, explaining language choices, responding to learner errors, and selecting examples that illustrate both regularity and variation. A teacher who has practiced inquiry-oriented analysis is better positioned to scaffold learners through comparable processes. Accordingly, the study examines not only measurable development in linguistic thinking but also participants' reported shifts in pedagogical orientation and professional self-understanding.

The Egyptian FFL teacher education context provides a meaningful site for this investigation. Digital transformation initiatives and institutional learning platforms have expanded access to technology, but there remains a need for models that connect this access to disciplined language inquiry and classroom pedagogy. In many contexts, pre-service teachers report confidence with everyday technologies while feeling less prepared to integrate digital tools in ways that transform subject-specific instruction. For FFL teacher education, this gap is visible when technology is used for presentation or communication but not for evidence-based language investigation.

The study therefore contributes to language teacher education in three ways. First, it offers an intervention model that operationalizes digital induction through concordance-based tasks and structured reflection. Second, it provides mixed-methods evidence on the development of linguistic thinking among pre-service FFL teachers. Third, it clarifies the affective and identity-related processes through which participants move from initial dependence on teacher-provided rules toward a more inquiry-oriented view of language teaching. The study is deliberately cautious in its claims: the design can demonstrate substantial within-group development and explain participants' experiences, but it cannot fully isolate the intervention from maturation, testing, or novelty effects.

Research Questions

1. To what extent do pre-service FFL teachers' linguistic thinking scores change after participation in a structured digital inductive model?
2. How do participants describe the model's influence on their pedagogical orientation and anticipated teaching practice?
3. What affordances and constraints do participants report while implementing corpus-supported inductive inquiry?

Literature Review

Digital Pedagogical Competence Beyond Tool Familiarity

Research on educational technology has repeatedly shown that pre-service teachers' everyday digital confidence does not automatically translate into transformative classroom practice. The TPACK framework conceptualizes effective technology integration as the intersection of technological, pedagogical, and content knowledge, while DigCompEdu specifies domains of educators' digital

competence such as professional engagement, digital resources, teaching and learning, assessment, learner empowerment, and learner facilitation (Mishra & Koehler, 2006; Redecker, 2017). Together, these frameworks shift attention from tool operation to pedagogical orchestration: the teacher's capacity to select and sequence digital resources because they serve specific disciplinary and learning goals.

This distinction matters for language teacher education. A learning management system, online dictionary, corpus, or concordancer can be used superficially, but such tools become pedagogically powerful only when embedded in tasks that require learners to compare examples, make interpretive decisions, and justify conclusions. Studies of digital competence development among pre-service teachers suggest that technology-rich courses are most valuable when they require active design, collaboration, and reflection rather than isolated technical training (Reisoglu & Cebi, 2020; Tondeur et al., 2012). Recent work in language teacher education similarly emphasizes the need to prepare teachers for digital environments through situated pedagogical practice rather than generic ICT exposure (Hughes et al., 2023; Ma et al., 2024).

Inductive Learning, Corpus Pedagogy, and Linguistic Thinking

Inductive learning positions learners as active constructors of knowledge who infer regularities from examples. In language learning, this orientation is particularly relevant because grammatical and lexical patterns are seldom fully captured by simplified rules. Corpora and concordancers create a digital environment in which learners can inspect multiple authentic instances of a linguistic phenomenon, notice distributional tendencies, identify exceptions, and connect form to use. This is the core rationale of data-driven learning (DDL), originally associated with the learner-as-researcher tradition and later expanded into diverse forms of corpus pedagogy (Boulton & Cobb, 2017; Johns, 1991).

The empirical literature on DDL suggests that corpus use can support language learning when tasks are carefully scaffolded and aligned with learners' proficiency and goals. Boulton and Cobb's meta-analysis reported positive overall effects for corpus use in language learning, but the literature also warns that unguided corpus exploration can overwhelm learners. For pre-service teachers, this tension is pedagogically useful. They need to experience both the power and the complexity of authentic language evidence so that they can later design realistic scaffolds for their own learners.

In this study, linguistic thinking refers to three related capacities: pattern recognition, hypothesis formulation, and metalinguistic reflection. Pattern recognition involves identifying regularities in authentic examples. Hypothesis formulation involves generating provisional explanations about why a pattern occurs and where it may not apply. Metalinguistic reflection involves articulating how linguistic evidence supports a claim about form, meaning, function, or context. These capacities are central to evidence-based grammar pedagogy and are directly compatible with inductive DDL tasks.

Pedagogical Identity and the Teacher-as-Guide

The study also attends to pedagogical identity because pre-service teachers' future classroom practice is influenced not only by what they know but also by how they understand the teacher's role. In deductive traditions, the teacher is often positioned as the authority who presents rules and verifies correct answers. In inquiry-oriented traditions, the teacher becomes a designer of tasks, a guide to evidence, and a facilitator of disciplined explanation. This shift is not simply technical; it is affective and professional. Participants who are accustomed to receiving rules may initially experience inductive inquiry as uncertain, inefficient, or cognitively demanding.

For this reason, a digital inductive model must be understood as both a cognitive intervention and a teacher-identity experience. If pre-service teachers learn through guided inquiry, they may become more capable of imagining similar learning for their future students. Conversely, if the experience is poorly scaffolded, it may reinforce the belief that corpus work is too difficult for classroom use. A sequential explanatory mixed-methods design is therefore appropriate because the quantitative phase can measure observed change in analytic performance, while the qualitative phase can explain how participants experienced the transition from rule dependency toward inquiry-based teaching.

Theoretical Framework

The framework conceptualizes the digital inductive model as an intersection of three domains. The first domain is digital pedagogical competence, which emphasizes purposeful selection and orchestration of technology for content-specific learning. The second is inductive and data-driven language inquiry, which provides the mechanism of pattern discovery, evidence-based generalization, and revision of hypotheses. The third is constructivist and socio-cultural learning, which frames knowledge construction as mediated through tools, dialogue, and scaffolded participation.

Figure 1 presents the logic of the framework. Authentic corpus evidence and concordance tasks are not treated as neutral add-ons; they are the mediating resources through which learners practice linguistic inquiry. Collaboration and instructor facilitation are included because corpus work becomes pedagogically meaningful when participants compare interpretations, challenge premature generalizations, and refine explanations. The expected outcomes are improved linguistic thinking and a more facilitative pedagogical identity.

Analytic Expectations

The quantitative hypothesis was limited to the construct measured by the LAT: participants would show statistically significant improvement in total linguistic thinking and in the three LAT subcomponents after the intervention. The qualitative expectations were explanatory rather than statistical: participants' accounts were expected to reveal how affordances such as authenticity, collaboration, and learner agency supported development, and how constraints such as technical friction, cognitive load, and prior deductive habits shaped implementation. This formulation aligns the claims with the actual data

sources and avoids treating qualitative perceptions as if they had been measured through a separate quantitative instrument.

Methods

Design

The study used a sequential explanatory mixed-methods design. Quantitative data were collected and analyzed first through a single-group pre-test/post-test design. Qualitative data were then used to explain and contextualize the observed quantitative change. This design was suitable because the research question concerned both measurable development in linguistic thinking and participants' interpretations of the learning process. However, the absence of a control or comparison group means that the quantitative findings should be interpreted as evidence of within-cohort change rather than as definitive evidence that the intervention alone caused the gains.

The manuscript therefore uses cautious causal language. Terms such as contributed to, was associated with, and supported are preferred over proved, demonstrated conclusively, or caused. This wording is consistent with the design and makes the study more defensible for peer review.

Participants and Setting

Participants were 56 third-year pre-service FFL student-teachers enrolled in an FFL didactics and pedagogy course at a public university in Beni-Suef, Egypt. Inclusion criteria were prior coursework in French grammar and pedagogy and no previous formal training in corpus-based instruction or data-driven learning. All participants completed the pre-test; one participant did not complete the post-test, yielding a matched quantitative sample of 55 for paired analyses. Participation was voluntary, and informed consent was obtained from all participants.

The study protocol was approved by the Faculty of Education Research Ethics Board, Beni-Suef University (Approval Reference No. BSU-FoE-003-01-3-2025), in accordance with the ethical standards governing research involving human participants and the 1964 Declaration of Helsinki and its later amendments. For the qualitative phase, a purposive subsample of 17 participants was selected to represent variation in baseline performance and engagement: 10 completed weekly reflective journals and 7 participated in a focus group.

Intervention: The Digital Inductive Model

The intervention was embedded in the regular course over ten weeks, with one 90-minute session per week. Each session focused on a target French grammatical or lexical phenomenon. Participants worked in small groups with a web-based concordancer and a purpose-built pedagogical corpus of authentic French texts. They searched for examples, examined concordance lines, categorized patterns,

discussed exceptions, wrote provisional generalizations, and revised those generalizations after peer and instructor feedback. Moodle was used for sharing group outputs and reflective responses.

The intervention deliberately sequenced support. Early sessions provided explicit demonstrations of concordance searching and pattern notation. Middle sessions shifted responsibility toward group-led analysis. Later sessions required participants to design brief pedagogical explanations or activities based on the patterns they identified. Table 1 summarizes the instructional architecture.

Table 1
Ten-week architecture of the digital inductive intervention

Phase	Weeks	Instructional focus	Participant output
Orientation and modeling	1–2	Introduction to concordance searching, corpus navigation, and evidence notation.	Annotated examples and guided pattern logs.
Guided pattern discovery	3–5	Collaborative analysis of target grammar and lexis through authentic examples.	Group generalizations with supporting concordance evidence.
Hypothesis testing	6–8	Attention to counterexamples, contextual constraints, and revision of rules.	Revised hypotheses and metalinguistic explanations.
Pedagogical transfer	9–10	Design of mini teaching applications and reflection on classroom feasibility.	Short inductive teaching tasks and reflective journals.

Instruments and Data Sources

The quantitative instrument was a researcher-developed Linguistic Analysis Task (LAT), administered immediately before and after the intervention. The task assessed three dimensions of linguistic thinking: pattern recognition, hypothesis formulation, and metalinguistic reflection. Items required participants to analyze authentic French examples and justify their interpretations from textual evidence. Content validity was established through review by three senior specialists in FFL pedagogy. Because the LAT was researcher-developed and context-specific, it should be treated as a performance measure rather than as a standardized scale.

Reflective journals provided the first qualitative source. Ten participants completed weekly digital journals documenting their learning processes, changing views of grammar instruction, and perceived affordances or challenges of the concordance-based tasks. The second qualitative source was a final semi-structured focus group with seven participants. The focus group lasted approximately 60 minutes, was audio-recorded, and was transcribed verbatim. Prompts explored perceived effectiveness, emotional response, collaboration, anticipated classroom transfer, and implementation constraints.

The LAT scoring framework is presented in Table 2.

Table 2
LAT scoring framework

Dimension	Maximum score	Core evidence of performance
Pattern recognition	5	Identifies recurrent grammatical or lexical patterns in authentic examples and distinguishes them from isolated instances.
Hypothesis formulation	5	Develops evidence-based provisional explanations and revises them when examples or counterexamples require refinement.
Metalinguistic reflection	10	Explains form-meaning relationships, contextual constraints, and pedagogical implications using explicit language evidence.
Total linguistic thinking	20	Composite score across the three dimensions.

Data Analysis

Quantitative data were analyzed using paired-samples t-tests on the matched sample (N = 55). Descriptive statistics were reported for pre-test and post-test scores. Statistical significance was set at $p < .05$, and effect sizes were reported using Cohen's d for paired samples. Because very large effects can occur in pre-test/post-test language education studies for several reasons, interpretation was situated within the design limitations and supported by qualitative evidence rather than treated as conclusive causal proof.

Qualitative data were analyzed thematically following Braun and Clarke's six-phase approach: familiarization, initial coding, theme development, theme review, theme definition, and narrative reporting. To strengthen dependability, a second researcher independently coded 20% of the dataset; disagreements were discussed until consensus was reached. The analysis prioritized themes that explained the quantitative pattern and addressed the second and third research questions. Integration occurred through narrative weaving and a joint display connecting quantitative findings with qualitative explanations.

Results

Quantitative Findings: Change in Linguistic Thinking

The matched quantitative sample included 55 participants. Table 3 presents the pre-test and post-test descriptive statistics. The total LAT score increased from $M = 9.14$ ($SD = 1.66$) to $M = 16.98$ ($SD = 1.90$). Gains were also observed across pattern recognition, hypothesis formulation, and metalinguistic reflection. To maintain consistency with the paired analysis, the table reports the matched sample size for both time points.

Table 3
LAT descriptive statistics for the matched sample (N = 55)

Measure	Pre-test M	Pre-test SD	Post-test M	Post-test SD
Pattern recognition (max = 5)	2.25	0.78	4.25	0.78
Hypothesis formulation (max = 5)	2.20	0.82	4.20	0.82
Metalinguistic reflection (max = 10)	4.69	1.06	8.53	1.05
Total linguistic thinking (max = 20)	9.14	1.66	16.98	1.90

Paired-samples t-tests confirmed statistically significant pre-post gains across the total LAT and all subcomponents (Table 4, Fig. 2). The total change was large, $t(54) = 21.43$, $p < .001$, $d = 2.89$. Pattern recognition and hypothesis formulation showed comparable gains, and metalinguistic reflection also improved substantially. Figure 3 presents the same pattern as percentage of maximum score to support visual comparison across subscales with different score ranges.

Table 4
Paired-samples t-tests comparing pre-test and post-test LAT scores (N = 55)

Comparison	Mean difference	SD difference	t	p	Cohen's d
Pattern recognition	2.00	0.82	17.26	< .001	2.33
Hypothesis formulation	2.00	0.82	17.26	< .001	2.33
Metalinguistic reflection	3.84	1.67	16.04	< .001	2.16
Total linguistic thinking	7.84	2.71	21.43	< .001	2.89

Qualitative Findings: Explaining the Observed Change

The qualitative analysis identified a developmental trajectory that explains how participants experienced the intervention. Four themes were central: initial uncertainty, cognitive engagement through authentic evidence, pedagogical reorientation, and emerging professional identity. These themes are summarized in Table 5 and represented in Fig. 2.

The first theme, initial uncertainty and technical friction, captured participants' early difficulty with the unfamiliarity of concordance work. Several participants described wanting the instructor to provide the rule directly. This response is important because it shows that the intervention challenged prior habits of learning. The model did not remove uncertainty; rather, it scaffolded participants through it.

The second theme, cognitive engagement through authentic evidence, was the most prominent explanation for quantitative improvement. Participants described the corpus as a space where language became investigable. Exceptions, which had previously been experienced as confusing, began to

function as clues for refining hypotheses. Collaboration supported this process by forcing groups to compare evidence and defend interpretations rather than accept a single answer prematurely.

The third theme, pedagogical reorientation, indicated a shift in how participants imagined grammar teaching. They increasingly described teaching as facilitation of discovery rather than transmission of rules. However, this shift was not naïve: participants also recognized that younger or lower-proficiency learners would require careful scaffolding, shorter searches, and curated examples.

The fourth theme, emerging professional identity, linked analytic confidence to future teaching. Participants reported feeling more capable of investigating unfamiliar language questions independently. This confidence supported a more facilitative identity in which the teacher guides learners toward evidence-based explanations.

Table 5
Summary of thematic findings

Theme	Core finding	Illustrative quotation
Initial uncertainty	Participants initially experienced concordance work as unfamiliar, technically demanding, and less secure than receiving a rule.	"At first, I wanted the teacher to give me the answer."
Cognitive engagement	Authentic examples and peer discussion encouraged pattern noticing, hypothesis testing, and durable understanding.	"The exceptions became clues, not just mistakes."
Pedagogical reorientation	Participants began to value guided discovery and learner-centered explanation while recognizing the need for scaffolding.	"This approach treats students like capable thinkers."
Emerging professional identity	Participants described themselves less as rule transmitters and more as guides who help learners investigate language evidence.	"I feel more like a language coach or guide now."

Integration of Quantitative and Qualitative Findings

The integrated findings show that the quantitative gains and qualitative trajectory are mutually explanatory. The LAT results indicate what changed: participants became more successful at identifying patterns, constructing evidence-based hypotheses, and reflecting metalinguistically. The qualitative results explain how that change was experienced: participants moved through uncertainty into sustained inquiry as authentic evidence and peer dialogue made language analysis visible and discussable. Table 6 presents the joint display linking research questions, findings, and interpretations.

Table 6
Joint display integrating quantitative and qualitative results

Research question	Quantitative finding	Qualitative explanation	Integrated interpretation
RQ1: Linguistic thinking	Total LAT and subcomponent scores improved significantly from pre-test to post-test.	Participants described repeated practice in noticing, comparing, and revising explanations.	The model was associated with substantial within-cohort development in analytic language reasoning.
RQ2: Pedagogical orientation	Not measured quantitatively.	Participants reported movement from rule-first teaching toward guided discovery and facilitation.	The intervention appears to have functioned as a pedagogical identity experience, not only a skills exercise.
RQ3: Affordances and constraints	Not measured quantitatively.	Authenticity and collaboration supported inquiry; technical friction and cognitive load initially constrained it.	Implementation should include technical orientation, curated examples, and progressive release of responsibility.

Discussion

Corpus-supported induction as a route to linguistic thinking

This study examined whether a structured digital inductive model could enhance pre-service French as a Foreign Language teachers' linguistic thinking and reshape how they understand their future pedagogical role. Taken together, the quantitative and qualitative findings provide convergent evidence that corpus-supported induction can function as more than a technological enrichment activity. When carefully scaffolded, it becomes a disciplined route into linguistic inquiry, enabling pre-service teachers to move from rule reception toward evidence-based analysis, hypothesis generation, and metalinguistic explanation.

The substantial improvement in Linguistic Analysis Task performance supports the central hypothesis that engagement with the digital inductive model would improve linguistic thinking. The gains across pattern recognition, hypothesis formulation, and metalinguistic reflection suggest that participants did not merely become more familiar with a specific digital tool; rather, they developed a more systematic way of approaching language evidence. This finding is consistent with the logic of data-driven learning, which positions learners as active investigators of authentic language rather than passive recipients of pre-formulated grammatical rules (Johns, 1991; Boulton & Cobb, 2017). In the present study, the concordancer did not operate as a neutral technological add-on. It provided repeated encounters with authentic examples, variation, frequency, counterexamples, and contextualized usage. These encounters

appear to have supported the cognitive operations targeted by the intervention: noticing patterns, comparing evidence, revising provisional explanations, and justifying claims about language use.

The qualitative findings clarify why the improvement occurred. Participants described a gradual change in how they approached linguistic problems. At the beginning, many expected the instructor to confirm the correct rule before they felt secure. By the end of the intervention, they increasingly framed language questions as investigable problems. This shift is pedagogically significant because linguistic thinking is not reducible to correct answers; it involves the ability to reason from evidence, tolerate ambiguity, and refine explanations when confronted with exceptions. The participants' descriptions of "detective work," collaborative comparison, and evidence-based explanation therefore provide process-level support for the quantitative gains. The intervention appears to have strengthened not only what participants knew about French forms but also how they learned to know: through inquiry, comparison, and principled interpretation.

This interpretation corroborates previous corpus-pedagogy research showing that corpus consultation can support noticing, discovery learning, and more durable understanding when learners are adequately trained and guided (Kennedy & Miceli, 2001; O'Keeffe et al., 2007; Boulton & Cobb, 2017). However, the contribution of the present study lies in its focus on pre-service teachers rather than language learners alone. For future teachers, corpus-supported induction has a dual function. It develops disciplinary understanding of language as patterned, variable, and evidence-sensitive, while also modeling a pedagogy they can later adapt for their own classrooms. In this sense, the model did not simply teach participants how to use a concordancer. It apprenticed them into a form of pedagogical reasoning in which language claims require evidence and teaching explanations emerge from guided inquiry.

The findings therefore support the first hypothesis directly and strongly: the digital inductive model was associated with significant gains in linguistic thinking. Nevertheless, the result should be interpreted with appropriate caution because the study used a single-group pre-test/post-test design. Without a comparison group, the observed gains cannot be attributed exclusively to the intervention with the same confidence as in a randomized or quasi-experimental design. Yet the convergence between the quantitative improvement and the qualitative accounts strengthens the plausibility of the interpretation. The qualitative data show that participants experienced precisely the forms of inquiry that the intervention was designed to cultivate, which makes the statistical improvement theoretically coherent rather than merely numerical.

From digital skills to digital pedagogical orchestration

A second major finding concerns the distinction between digital tool familiarity and digital pedagogical competence. Participants' early difficulty with the concordancer indicates that being comfortable with general digital technologies does not automatically prepare pre-service teachers to use specialized tools for disciplinary learning. The concordancer required them to formulate search questions, interpret concordance lines, distinguish salient from incidental patterns, identify counterexamples, and transform

raw evidence into a pedagogically meaningful explanation. These demands required the integration of technological knowledge, content knowledge, and pedagogical knowledge, which is central to the TPACK framework (Mishra & Koehler, 2006; Koehler et al., 2013).

The findings therefore support the view that effective technology integration in teacher education depends on pedagogical orchestration rather than tool exposure. A concordancer can remain a technical novelty if it is introduced only through demonstration. It becomes pedagogically powerful when embedded in a sequence of tasks that moves learners from guided exploration to collaborative interpretation, from pattern detection to hypothesis testing, and from linguistic analysis to teaching design. This progression also aligns with DigCompEdu's emphasis on digitally supported teaching, learning, learner empowerment, and reflective professional competence (Redecker & Punie, 2017; Reisoğlu & Çebi, 2020). In the present study, participants did not simply acquire operational digital skills. They learned to use a digital resource in the service of disciplinary reasoning and future instructional decision-making.

This point implicitly addresses the study's concern with pedagogical orientation. The intervention appears to have shifted participants' understanding of what technology is for in language teaching. At the outset, some participants treated the tool as an obstacle or as a complicated way to reach an answer that the teacher could have supplied more directly. Over time, they began to understand the tool as a mediator of inquiry. This shift matters because many technology-integration initiatives assume that access, training, and exposure are sufficient. The present findings suggest otherwise. Digital pedagogical competence emerges when teachers experience how technology changes the structure of learning activity: who asks the question, where evidence comes from, how interpretations are tested, and how learners participate in constructing knowledge.

The findings also verify the hypothesis that the model would provide distinctive affordances compared with traditional rule-first instruction, although this verification is qualitative rather than statistical. Participants repeatedly identified authenticity and collaboration as the model's most important affordances. Authenticity gave them access to language as actually used rather than language as simplified for textbook presentation. Collaboration allowed them to compare interpretations, challenge premature conclusions, and negotiate explanations. These affordances are central to DDL because corpus evidence is rarely self-explanatory. Learners require interaction, guidance, and reflective prompts to transform examples into understanding. Thus, the model's effectiveness depended on the interaction between digital evidence and social scaffolding, not on the corpus alone.

Pedagogical identity as an outcome of learning design

One of the most important qualitative contributions of the study is the link it establishes between learning design and pedagogical identity. Participants' descriptions suggest that the intervention affected how they imagined themselves as future teachers. They increasingly described the teacher not as a transmitter of rules or a final authority on correctness, but as a guide, coach, or facilitator of inquiry. This finding supports the hypothesis that participation in the digital inductive model would encourage a

shift from deductive, teacher-centered orientations toward more inquiry-based and learner-centered pedagogical beliefs.

The significance of this finding lies in the coherence between method and message. Teacher education programs frequently encourage learner-centered pedagogy while relying heavily on transmissive instruction. Such inconsistency can limit the credibility and transferability of pedagogical reform. In the present intervention, participants experienced the kind of learning they were being invited to design: guided discovery, evidence-based reasoning, peer negotiation, and reflective explanation. The resulting identity shift is therefore not surprising. It emerged from participation in a different pedagogical culture. Participants learned, through experience, that uncertainty can be productive, that learners can generate valid explanations when supported, and that the teacher's role can involve structuring inquiry rather than pre-empting it.

This finding resonates with teacher-identity research emphasizing that professional identity develops through situated participation, reflection, and meaningful pedagogical experience rather than through abstract knowledge alone (Beauchamp & Thomas, 2009; Korthagen, 2017). In the present study, identity change was not treated as a separate affective outcome; it was embedded in the learning process. As participants became more capable of analyzing language evidence, they also became more willing to imagine their future learners as capable of inquiry. This is an important contribution because it suggests that cognitive development and pedagogical identity development can be mutually reinforcing. Learning to think linguistically may help future teachers learn to teach linguistically.

At the same time, the findings caution against an overly idealized view of discovery learning. Participants' development was not immediate or frictionless. The qualitative data indicate that anxiety, cognitive load, technical frustration, and dependence on teacher confirmation were genuine constraints. These findings support the hypothesis that participants would experience implementation challenges connected to cognitive demand and prior deductive habits. However, the evidence also qualifies that hypothesis. The constraints did not prevent learning; rather, they became part of the developmental trajectory when adequate scaffolding was available. Initial anxiety became manageable as participants gained procedural familiarity, collaborative support, and repeated opportunities to test interpretations.

This has an important theoretical implication. Induction should not be equated with unguided discovery. The model's effectiveness appears to have depended on structured inquiry: curated corpora, clear task prompts, peer discussion, instructor facilitation, and staged movement from guided to more independent analysis. These scaffolds did not dilute inductive learning. They made it cognitively and emotionally accessible. This interpretation is consistent with socio-cultural accounts of learning, in which mediation and interaction support learners' movement from dependence toward greater independence (Vygotsky, 1978). In the present study, the concordancer, peer group, task sheet, instructor prompts, and reflective journals functioned as mediating resources. Together, they helped participants transform uncertainty into inquiry and inquiry into professional confidence.

Reframing the role of authenticity and collaboration

The study also contributes to understanding the conditions under which authentic corpus evidence becomes pedagogically useful. Authenticity was one of the strongest affordances reported by participants, but the findings show that authenticity alone is insufficient. Authentic data can overwhelm learners if it is introduced without guidance. The participants initially encountered complexity, exceptions, and conflicting examples as sources of confusion. Over time, however, these same features became resources for deeper understanding. This transformation depended on the design of tasks that encouraged comparison, explanation, and revision.

The pedagogical value of authenticity therefore lies not simply in exposing learners to real language, but in helping them develop disciplined ways of reading real language. Corpus evidence is valuable because it reveals distribution, variation, collocation, patterning, and contextual use. Yet these features become visible only when learners have analytic questions and interpretive procedures. The present study suggests that pre-service teachers can develop these procedures when they are guided to search purposefully, document evidence, discuss anomalies, and revise generalizations. In this way, authenticity becomes a route to linguistic thinking rather than a source of unstructured complexity.

Collaboration played a similarly important role. Participants' accounts suggest that peer interaction helped them move beyond individual uncertainty. Through group discussion, they could compare interpretations, challenge weak explanations, and recognize alternative readings of the same evidence. This finding reinforces the socio-cultural dimension of the theoretical framework. Corpus inquiry was not an isolated cognitive activity; it was socially mediated. The model created a learning environment in which evidence, dialogue, and instructor facilitation worked together. This explains why participants' growth was both analytic and pedagogical. They experienced how collaborative reasoning can support language learning, and this experience informed how they imagined their future teaching.

Implications for FFL teacher education

The findings have several implications for French as a Foreign Language teacher education in Egypt and similar contexts where digital transformation is often associated with access to tools, platforms, and infrastructure. The study suggests that teacher education programs should move beyond technical digital literacy and integrate corpus-supported induction into methodology, grammar, lexis, and discourse-analysis courses. Such integration would allow pre-service teachers to develop linguistic thinking through repeated engagement with authentic evidence while simultaneously learning how inquiry-based tasks can be designed for learners.

First, corpus-supported induction should be treated as a pedagogical model rather than as an optional technical skill. If introduced only as a software exercise, concordancing may remain peripheral to teachers' professional practice. If embedded within didactics courses, however, it can help future teachers understand the relationship between language evidence, explanation, and instructional design. Second, teacher education programs should begin with small, curated pedagogical corpora before expecting students to use large open corpora independently. Curated corpora reduce unnecessary cognitive load while preserving the central benefits of authenticity and pattern discovery. Third,

assessment should value evidence-based reasoning, not only rule recall. Tasks such as the LAT are important because they require participants to justify explanations through examples rather than reproduce definitions.

Fourth, pre-service teachers should be required to translate corpus findings into classroom-ready activities. This step is crucial for transfer. It helps participants move from “I can analyze language” to “I can help learners analyze language.” In practical terms, methodology courses could ask student-teachers to design guided concordance tasks, prepare learner-friendly evidence sheets, anticipate misconceptions, and formulate scaffolding questions. Such activities would strengthen the link between linguistic analysis and pedagogical planning.

For curriculum planners, the findings indicate that the model is feasible because it can be embedded in existing teacher-education courses without requiring full program redesign. A ten-week sequence can be aligned with grammar instruction, vocabulary development, discourse analysis, error analysis, or materials design. The essential condition is not the sophistication of the software but the quality of the inquiry sequence. The intervention must help participants ask searchable questions, interpret patterns, discuss exceptions, formulate claims, and reflect on classroom transfer. Technology becomes transformative only when it is organized around these pedagogical actions.

Theoretical contribution and qualified interpretation

Theoretically, the study strengthens the argument that linguistic thinking develops at the intersection of authentic evidence, scaffolded inquiry, digital mediation, and professional reflection. The results support the integrated framework of the study: digital pedagogical competence provides the capacity to use tools purposefully; inductive and data-driven learning provides the mechanism of pattern discovery and hypothesis revision; socio-cultural learning explains the role of collaboration and instructor mediation; and teacher-identity development explains how changed learning experiences can reshape future pedagogical orientations.

The study also offers a more nuanced interpretation of digital transformation in language teacher education. Digital transformation should not be measured only by whether teachers can access or operate tools. A more meaningful indicator is whether they can use digital environments to create forms of learning that would be difficult to achieve through traditional rule transmission. In this study, the digital inductive model created such a form of learning. It gave pre-service teachers access to authentic language evidence, required them to reason from examples, encouraged collaborative interpretation, and connected analysis to future teaching. This combination explains why the intervention affected both linguistic thinking and pedagogical identity.

At the same time, the study’s claims must remain appropriately bounded. The design supports strong evidence of pre-post improvement, but it does not prove superiority over all alternative forms of instruction. The qualitative findings provide rich evidence of perceived identity transformation,

affordances, and constraints, but these perceptions need to be examined longitudinally to determine whether they translate into sustained classroom practice. Future studies should include control or comparison groups, delayed post-tests, classroom observation during practicum, and more extensive validation of the LAT. Further research could also compare different levels of scaffolding, different types of corpora, and different proficiency groups to determine which conditions best support transfer.

Despite these limitations, the findings offer a coherent and practically useful conclusion. The digital inductive model supported measurable growth in linguistic thinking, encouraged a shift toward inquiry-oriented pedagogical beliefs, and revealed both the affordances and constraints of corpus-supported teacher education. The hypotheses were therefore supported in a differentiated manner: the improvement in linguistic thinking was directly supported by quantitative evidence, while the shifts in pedagogical orientation, perceived affordances, and implementation constraints were supported by qualitative evidence. The overall pattern indicates that corpus-supported induction can be a productive pathway for preparing pre-service FFL teachers who are not only digitally competent but also analytically confident, pedagogically reflective, and capable of designing language learning as guided inquiry.

Limitations and Directions for Future Research

The study has several limitations that should be addressed in future research. First, the quantitative phase used a single-group pre-test/post-test design. This design is appropriate for exploratory intervention evaluation but cannot rule out testing effects, maturation, novelty effects, or concurrent learning from the wider course. Future studies should use randomized, quasi-experimental, or active-comparator designs when feasible.

Second, the LAT was researcher-developed. Although expert review supported content validity, the instrument requires further psychometric development, including inter-rater reliability, internal consistency where appropriate, and evidence of sensitivity across proficiency levels. Future manuscripts should include sample items and a full scoring rubric as supplementary material.

Third, the study was conducted at one public university with pre-service FFL teachers, limiting transferability to other languages, institutions, and educational systems. Fourth, qualitative data came from a purposive subsample of 17 participants. The sample was adequate for explanatory depth, but additional interviews, classroom observations, and longitudinal follow-up would strengthen claims about identity development. Finally, the study measured pedagogical orientation and anticipated transfer rather than actual classroom implementation. Future research should examine whether pre-service teachers later use corpus-supported induction during practicum or early career teaching.

Conclusion

This study provides a methodologically bounded account of a digital inductive model for developing linguistic thinking among pre-service FFL teachers. Quantitative findings showed substantial within-cohort gains on a Linguistic Analysis Task after ten weeks of guided concordance-based inquiry.

Qualitative findings explained these gains as part of a developmental process in which participants moved from uncertainty and rule dependence toward authentic evidence use, collaborative reasoning, and a more facilitative view of teaching.

The central contribution is not that technology alone improves teacher preparation. Rather, the study shows how technology can become pedagogically consequential when embedded in an inductive sequence that requires learners to search, compare, hypothesize, explain, and reflect. In this model, the concordancer functions as a mediator of disciplined language inquiry, and collaboration functions as a scaffold for interpretation. For FFL teacher education, this offers a practical pathway for connecting digital competence with analytic content knowledge and pedagogical identity formation.

The interpretation remains cautious. The absence of a control group prevents definitive causal attribution, and the researcher-developed assessment requires further validation. Nevertheless, the convergence between the quantitative pattern and qualitative explanation supports the value of structured, corpus-supported induction as a promising approach in technology-enhanced language teacher education. Future controlled and longitudinal studies can build on this model to determine how durable the gains are and how effectively pre-service teachers transfer inquiry-oriented language analysis into classroom practice.

Declarations

Ethical Approval and Consent to Participate

The study protocol was approved by the Faculty of Education Research Ethics Board, Beni-Suef University (Approval Reference No. BSU-FoE-003-01-3-2025), in accordance with the ethical standards governing research involving human participants and the 1964 Declaration of Helsinki and its later amendments. Informed consent was obtained from all individual participants included in the study.

Consent for Publication

Not applicable.

Availability of Data and Materials

The datasets generated and analyzed during the current study are not publicly available because of participant privacy and ethical restrictions but may be made available from the corresponding author on reasonable request and subject to institutional approval.

Competing Interests

The authors declare that they have no competing interests.

Funding

No funds, grants, or other support were received for conducting this study or preparing the manuscript.

Authors' Contributions

M.A.M. and F.A. contributed to study conception and design. M.A.M. and F.A. contributed to material preparation, data collection, and analysis. M.A.M. drafted the manuscript. Both authors critically reviewed the manuscript, approved the final version, and agreed to be accountable for the accuracy and integrity of the work.

Acknowledgements

The authors acknowledge the institutional support of Beni-Suef University. The acknowledgement of potential APC support should be included only if confirmed by the relevant funding body at the time of submission.

Authors' Information

Mohamed Amin Mekheimer is Professor of Applied Linguistics (TEFL), Faculty of Education, Beni-Suef University, Beni-Suef, Egypt. Fatma Abdelaal is Associate Professor of Applied Linguistics (Teaching French as a Foreign Language), Faculty of Education, Beni-Suef University, Beni-Suef, Egypt.

References

1. Aleksieva L (2025) Preparing pre-service teachers for the digital transformation of education: Exploring university teacher educators' views and practical strategies. *Educ Sci* 15(4). Article 404. <https://doi.org/10.3390/educsci15040404>
2. Boulton A, Cobb T (2017) Corpus use in language learning: A meta-analysis. *Lang Learn* 67(2):348–393. <https://doi.org/10.1111/lang.12224>
3. Braun V, Clarke V (2006) Using thematic analysis in psychology. *Qualitative Res Psychol* 3(2):77–101. <https://doi.org/10.1191/1478088706qp063oa>
4. Creswell JW, Plano Clark VL (2018) *Designing and conducting mixed methods research*, 3rd edn. SAGE
5. Hughes SP, Corral-Robles S, Ortega-Martín JL (2023) Let's get digital: ICT training needs in pre-service language teaching. *Educ Sci* 13(12). Article 1238. <https://doi.org/10.3390/educsci13121238>

6. Johns T (1991) Should you be persuaded: Two examples of data-driven learning. In: Johns T, King P (eds) Classroom concordancing. Centre for English Language Studies, University of Birmingham, pp 1–16
7. Kharade K, Peese H (2014) Problem-based learning: A promising pathway for empowering pre-service teachers for ICT-mediated language teaching. *Policy Futures Educ* 12(2):262–272. <https://doi.org/10.2304/pfie.2014.12.2.262>
8. Ma Q, Lee HTH, Gao X, Chai C-S (2024) Learning by design: Enhancing online collaboration in developing pre-service TESOL teachers' TPACK for teaching with corpus technology. *Br J Edu Technol* 55(6):2639–2667. <https://doi.org/10.1111/bjet.13458>
9. Mishra P, Koehler MJ (2006) Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers Coll Record* 108(6):1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
10. Moorhouse BL, Wan Y, Wu C, Kohnke L, Ho TY, Kwong T (2024) Developing language teachers' professional generative AI competence: An intervention study in an initial language teacher education course. *System* 125:103399. <https://doi.org/10.1016/j.system.2024.103399>
11. O'Keeffe A, McCarthy M, Carter R (2007) From corpus to classroom: Language use and language teaching. Cambridge University Press
12. Plonsky L, Oswald FL (2014) How big is big? Interpreting effect sizes in L2 research. *Lang Learn* 64(4):878–912. <https://doi.org/10.1111/lang.12079>
13. Redecker C (2017) In: Punie Y (ed) European framework for the digital competence of educators: DigCompEdu. Publications Office of the European Union. <https://doi.org/10.2760/159770>
14. Reisoğlu I, Çebi A (2020) How can the digital competences of pre-service teachers be developed? Examining a case study through the lens of DigComp and DigCompEdu. *Comput Educ* 156:103940. <https://doi.org/10.1016/j.compedu.2020.103940>
15. Tondeur J, van Braak J, Sang G, Voogt J, Fisser P, Ottenbreit-Leftwich A (2012) Preparing pre-service teachers to integrate technology in education: A synthesis of qualitative evidence. *Comput Educ* 59(1):134–144. <https://doi.org/10.1016/j.compedu.2011.10.009>
16. Vygotsky LS (1978) *Mind in society: The development of higher psychological processes*. Harvard University Press

Figures

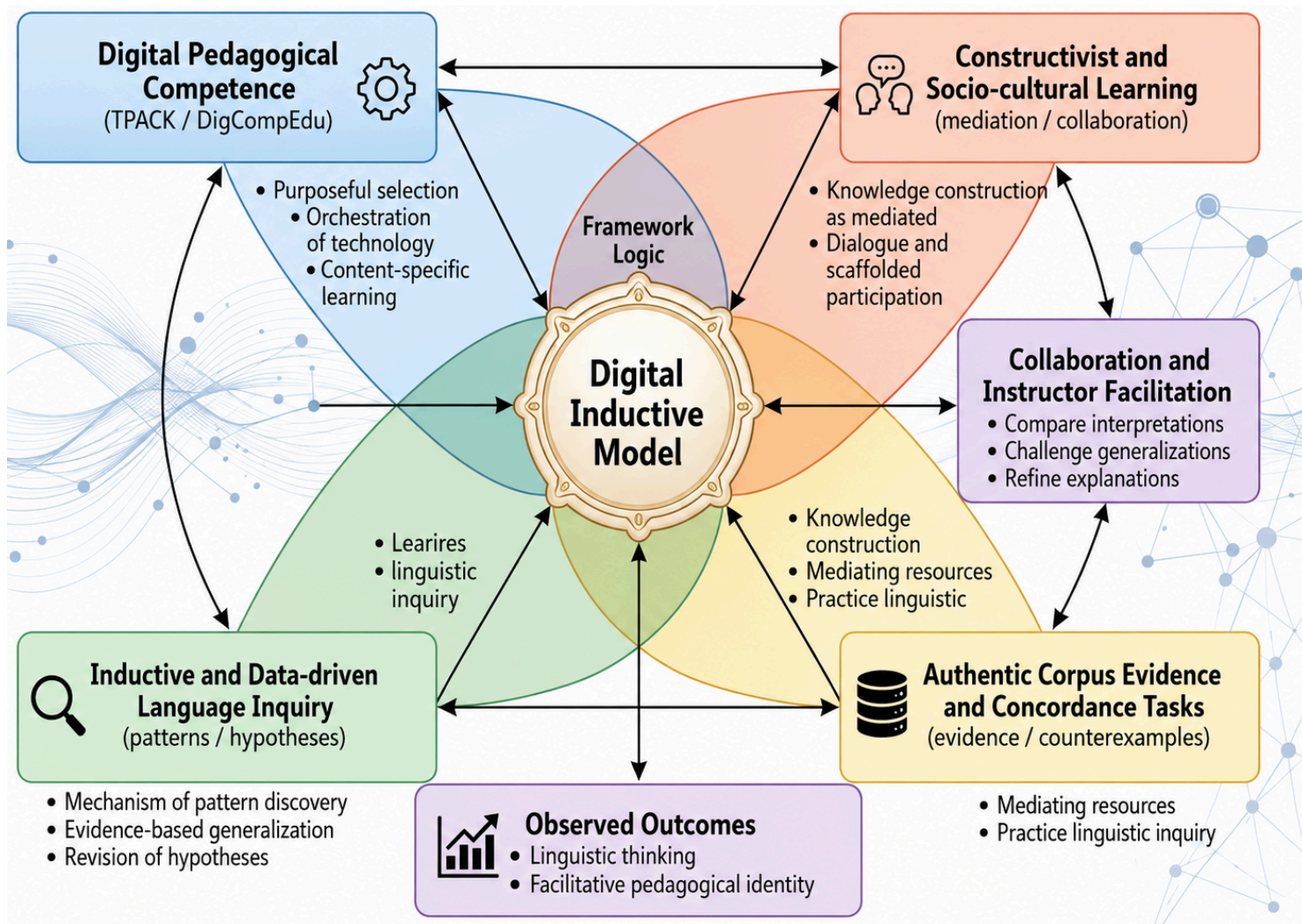


Figure 1

Conceptual framework of the digital inductive model.

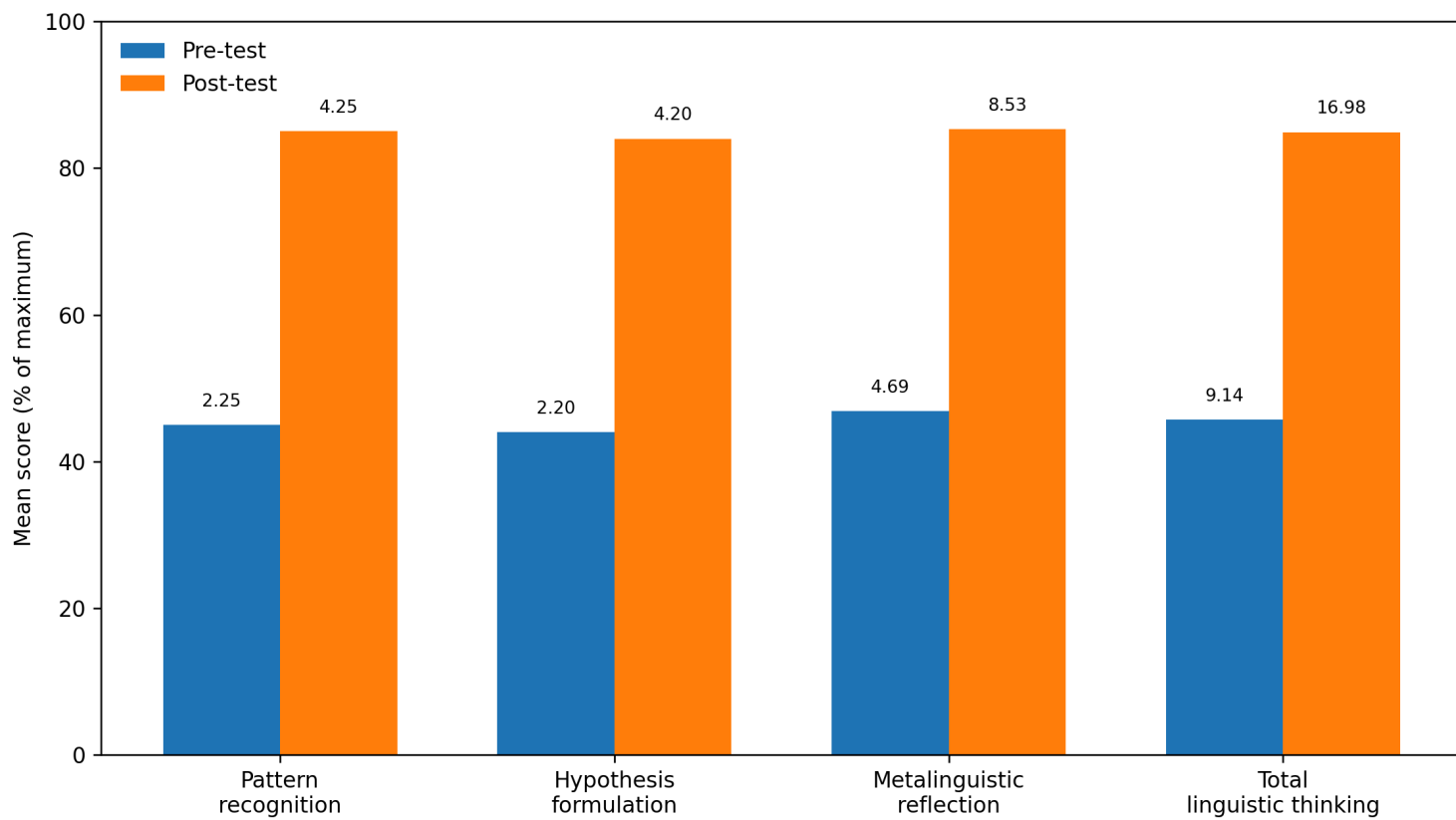


Figure 2

Observed pre-post change on the Linguistic Analysis Task. Scores are shown as percentages of the maximum possible score; raw means appear above bars.

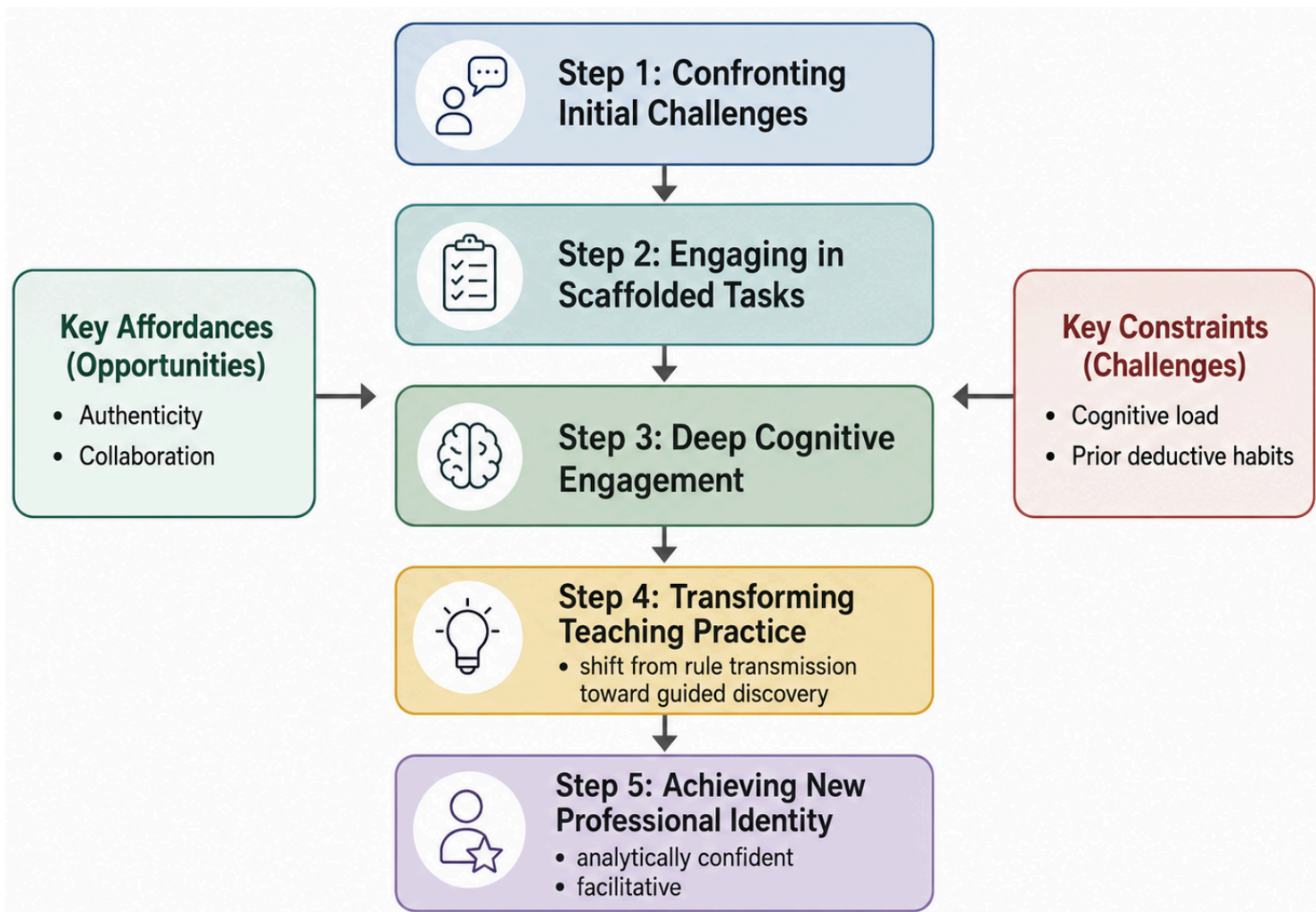


Figure 3

Qualitative process model of participants' reported experience.