Materiallar to’plami · Conference Proceedings

TIL, TARJIMA VA TA’LIM: GLOBAL KOMMUNIKATSIYA VA RAQAMLI INNOVATSIYALAR DAVRIDA YANGI YONDASHUVLAR

03 October 2026

UDK 378

TIL O‘QITISH METODIKASI VA RAQAMLI PEDAGOGIKA

Oliy ta’limda ingliz tilini mustaqil o‘rganish uchun raqamli kompetensiya sifatida korpus savodxonligini rivojlantirish

Munojat Sultonova*

Farg’ona davlat universiteti

✉ msultonova27@gmail.com

ANNOTATSIYA

Raqamli texnologiyalarning jadal rivojlanishi universitet talabalarining xorijiy tillarni o’rganish va lingvistik ma’lumotlarga kirish usullarini o’zgartirdi. Zamonaviy oliy ta’limda mustaqil ingliz tilini o’rganish talabalardan nafaqat lug’at va grammatikani bilishni, balki haqiqiy til dalillarini topish, baholash, talqin qilish va qo’llashni ham talab qiladi. Shu nuqtai nazardan, korpus savodxonligini muhim zamonaviy raqamli va akademik kompetensiya sifatida ko’rib chiqish mumkin. Ushbu maqolada korpus savodxonligi oliy ta’limdagi filologiya talabalari orasida mustaqil ingliz tilini o’rganishning tarkibiy qismi sifatida, xususan, Farg’ona davlat universiteti uchun ham o’rganiladi. Maqolada korpusga asoslangan o’qitish, ma’lumotlarga asoslangan o’qitish (DDL), o’quvchining mustaqil o’rganishi va mustaqil korpus konsultatsiyasi bo’yicha so’nggi tadqiqotlar ko’rib chiqiladi. Lug’at, kollokatsiya, grammatika va akademik yozuvni rivojlantirish uchun resurslar sifatida Zamonaviy Amerika ingliz tili korpusi (COCA), Britaniya milliy korpusi (BNC), AntConc, LancsBox va SKELLdan foydalanishga alohida e’tibor qaratilgan. So’nggi tizimli sharhlar shuni ko’rsatadiki, korpusga asoslangan o’qitish o’quvchining avtonomiyasi va ingliz tilini rivojlantirishga yordam berishi mumkin, shu bilan birga korpusning murakkabligi, texnik ko’nikmalar, korpus dalillarini talqin qilish va yetarli bo’lmagan tayyorgarlik bilan bog’liq muammolarni ham ochib beradi. Ko’rib chiqilgan adabiyotlarga asoslanib, maqolada korpus savodxonligining besh o’lchovli modeli taklif qilinadi: korpus xabardorligi, texnik qidiruv kompetensiyasi, korpus ma’lumotlarini talqin qilish, tanqidiy baholash va mustaqil qo’llash. Maqolada korpus savodxonligi ingliz tili va filologiya dasturlariga talabalarga tilni mustaqil o’rganish va o’zini o’zi boshqarish qobiliyatiga ega o’quvchilarga aylanish imkonini beruvchi raqamli akademik kompetensiya sifatida integratsiya qilinishi kerakligi ta’kidlangan.

Kalit so‘zlar

korpus savodxonligi, raqamli kompetensiya, mustaqil ta’lim, o‘quvchi avtonomiyasi, ma’lumotlarga asoslangan ta’lim, korpus lingvistikasi, COCA, BNC, AntConc, LancsBox, SKELL, oliy ta’lim, filologiya talabalari.

DOI: 10.5281/zenodo.23278544

Introduction

Digital transformation has changed the role of the university student from a passive recipient of information to an active participant in knowledge construction. This change is particularly important in foreign-language education. University students are increasingly expected to continue learning outside the classroom, identify their own language needs, use digital resources and evaluate information independently.

For Philology students, this expectation is especially significant. Their future professional activities may include teaching English, translation, academic writing, linguistic research and language assessment. Consequently, they need more than knowledge of linguistic rules. They need to understand how English is actually used in authentic contexts and how linguistic evidence can be investigated independently.

Corpus linguistics provides an important technological foundation for this process. A corpus is an electronically stored collection of authentic language that can be searched and analysed using digital tools. Instead of depending exclusively on a teacher, textbook or dictionary, students can investigate thousands of

examples and identify patterns themselves.

This approach is commonly associated with **Data-Driven Learning (DDL)**. DDL encourages learners to explore corpus data and develop their own understanding of language through observation and analysis. Recent research indicates that DDL is increasingly being reconsidered in relation to learner autonomy, digital pedagogy and independent learning.

Casal and Kessler (2024), introducing a special issue of *TESOL Quarterly*, identify four current directions in DDL research: selecting appropriate corpora, extending the boundaries of DDL, preparing language educators for corpus-based teaching, and developing stronger theoretical and methodological foundations. Their article appears in *TESOL Quarterly, 58*(3), 1027–1045.

Particularly important is the systematic review by Li, Noordin, Ismail and Cao (2025). The authors examined empirical research published between 2011 and 2024 and selected **44 studies** for their final review. Their findings indicate that corpus-based instruction can support the development of English skills and learner autonomy, while significant challenges include corpus complexity, the concentration of research on higher-level learners and instructional limitations. They recommend comprehensive training, simplified resources and tasks, personalised learning paths and greater resource provision.

Therefore, the central argument of this article is that **corpus literacy should be understood not merely as a linguistic research skill but as a modern digital competence for independent English language learning in higher education.**

The article addresses three questions:

1. What corpus skills do Philology students need for independent English learning?
2. How can corpus literacy contribute to learner autonomy and independent learning?
3. What barriers prevent students from using corpus tools independently?

Corpus Literacy as a Digital and Academic Competence

The concept of corpus literacy is not entirely new. Mukherjee's influential conceptualisation identified four major dimensions: understanding what a corpus is, understanding what can and cannot be done with a corpus, knowing how to analyse corpus data and knowing how to draw conclusions about language use from corpus evidence.

Farr and Leńko-Szymańska (2024) continue to use this conceptualisation in their discussion of corpus integration into teacher education. They explain corpus literacy as including:

"understanding what a corpus is"

as well as knowing what can and cannot be done with corpora, analysing corpus data and drawing conclusions about language use.

Their article proposes a broader framework in which corpus literacy is integrated with corpus-based language pedagogy and corpus-based reflective practice. The authors argue that corpus integration should not be limited to introducing teachers to corpus websites; rather, learners and future teachers need structured opportunities to develop the skills necessary to use corpus evidence. The article appears in *TESOL Quarterly, 58*(3), 1181–1192.

For higher education, corpus literacy can therefore be conceptualised as a **digital academic competence** consisting of five interconnected dimensions:

2.1. Corpus awareness

Students should understand:

- what a corpus is;
- how corpora differ from ordinary Internet searches;
- what kinds of language a corpus contains;
- how corpus size and composition influence results;
- the difference between general, specialised and learner corpora.

2.2. Technical-search competence

Students should know how to:

- search keywords;
- use phrase searches;
- examine concordance lines;
- identify collocations;
- compare frequencies;
- use filters;
- search different registers or genres;
- download or organise corpus results.

2.3. Data interpretation

Students should be able to move beyond simply collecting examples.

They should ask:

- What pattern occurs repeatedly?
- What words commonly occur together?
- In which contexts does this expression occur?
- Is the pattern characteristic of academic or informal English?

2.4. Critical evaluation

Students should understand that corpus evidence is not automatically an absolute linguistic rule.

They need to consider:

- corpus size;
- representativeness;
- genre;
- register;
- frequency;
- context;
- date of the data;
- limitations of the corpus.

2.5. Independent application

Finally, students should be able to use corpus evidence to solve their own language problems without continuous teacher intervention.

This final dimension connects corpus literacy directly to independent learning.

Corpus Tools for Independent English Learning

Different corpus tools can support different aspects of independent learning.

Corpus of Contemporary American English (COCA)

COCA provides a large collection of contemporary American English and allows learners to investigate vocabulary, grammar, collocation, register and phraseological patterns.

For example, a student writing:

make a research can investigate *research* and discover frequent patterns such as:

- conduct research;
- carry out research;
- undertake research;

- do research.

The student therefore does not simply receive a correction from a teacher. The student learns **how to investigate the correction**.

British National Corpus (BNC)

BNC can provide another perspective on English usage, particularly British English.

Students can compare expressions in COCA and BNC and investigate differences in:

- frequency;
- register;
- lexical preference;
- grammatical patterns.

Such comparisons are particularly useful for Philology students because they develop linguistic awareness rather than simple memorisation.

AntConc

AntConc is particularly useful when students want to analyse their own texts. Students can create a small personal learner corpus from:

-
- essays;
- research papers;
- assignments;
- translations;
- presentations.

They can then investigate:

- word frequency;
- concordance lines;
- clusters;
- keywords;
- collocations.

This creates a bridge between **corpus linguistics and reflective learning**. LancsBox provides tools for exploring collocations, frequency and relationships between words in corpus data.

Students can investigate an academic word such as:

Significant and examine expressions such as:

- significant increase;
- significant difference;
- statistically significant;
- significant impact.

The learner begins to understand vocabulary as a network of relationships rather than isolated dictionary definitions.

SKELL can be particularly useful for students who are beginning corpus consultation because it provides relatively accessible examples of words and collocations.

Students can investigate whether a combination such as:

powerful research

or

strong research

is more typical in English and examine examples before making their own decision.

Thus, different corpus tools can be integrated according to the student's proficiency level and learning objective.

Independent learning in higher education involves students taking increasing responsibility for identifying learning needs, selecting resources, monitoring progress and evaluating outcomes.

Corpus tools can operationalise these principles.A traditional learning process may be represented as:**Teacher → explanation → exercise → correction**A corpus-based independent learning process can instead be represented as:**Language problem → corpus search → evidence → interpretation → hypothesis → application → reflection**For example:**Step 1: Identify the problem**The student is unsure whether:*discuss about the problem*is correct.**Step 2: Search the corpus**The student searches:**discuss + nounStep 3: Examine evidence**The student observes examples such as:

- discuss the problem;
- discuss the issue;
- discuss the results.

Step 4: Formulate a conclusion

The learner concludes that *discuss* normally takes its object directly and does not require *about* in this construction. The first research question can therefore be answered through a competency framework.

Corpus competence	Student ability	Independent-learning application
Corpus awareness	Understand corpus principles	Select an appropriate resource
Search competence	Construct searches	Investigate language independently
Concordance literacy	Read concordance lines	Identify usage patterns
Collocation competence	Find word combinations	Improve vocabulary
Frequency literacy	Interpret frequency	Compare alternatives
Register awareness	Identify context/register	Improve academic writing
Critical evaluation	Evaluate evidence	Avoid overgeneralisation
Reflective competence	Record findings	Build personal learning resources
Transfer competence	Apply findings	Improve own English

The most important point is that corpus literacy is **multidimensional**. It combines technical, linguistic, analytical and metacognitive skills.

The second research question concerns the relationship between corpus literacy and independent learning. The reviewed literature suggests several mechanisms.

It increases access to authentic evidence Students can investigate real language instead of relying exclusively on simplified textbook examples. **It develops learner agency**Students can formulate and investigate their own linguistic questions.

It encourages discovery learning Students identify patterns rather than simply receiving rules. **It develops metalinguistic awareness**Students begin to notice:

- collocations;
- phraseology;
- register;
- grammatical patterns;
- semantic preferences.

It supports self-correction Students can investigate their own language problems before asking the teacher. **It develops research competence**Students learn to formulate questions, collect evidence and draw conclusions.These processes are consistent with the autonomy-oriented direction of contemporary DDL research. Casal and Kessler (2024) identify learner and teacher development, corpus selection and theoretical reflection as important current areas for DDL research.

The literature reviewed in this article demonstrates that corpus technology has considerable potential for higher education, but technology alone does not produce autonomous learning.This distinction is important. A student may have access to COCA but lack corpus literacy. A student may know AntConc but not know how to interpret frequency.A student may see hundreds of concordance lines but be unable to identify a meaningful pattern.

Access to technology is not equivalent to digital competence.

Corpus literacy represents the knowledge required to transform digital access into meaningful learning.

This conclusion is supported by Hua et al. (2024), whose participants had access to an independent corpus tool but fewer than 40% reported frequent use.

It is also consistent with Li et al. (2025), who conclude from 44 empirical studies that corpus-based instruction can promote learner autonomy but requires comprehensive training and appropriately simplified tasks.

Consequently, higher education institutions should not simply provide students with links to corpus websites. They should develop **curriculum-based corpus literacy**.The teacher's role also changes.The teacher initially acts as:**demonstrator → trainer → facilitator → adviser**while students gradually move from:**dependent learner → guided corpus user → independent corpus consultant → autonomous language investigator.**This gradual transfer of responsibility is particularly appropriate for university Philology students because their academic and professional development ultimately requires independent research and continuous learning.

Corpus literacy should be considered an important component of digital competence in contemporary higher education English language learning. It enables students to move beyond textbook-based language learning and investigate authentic language evidence independently.

The reviewed research demonstrates that corpus-based instruction and Data-Driven Learning can support English language development and learner autonomy. Li et al. (2025), based on a systematic review of 44 empirical studies, identify positive effects while also emphasising the challenges of corpus complexity, learner preparation and instructional design.

Recent research also demonstrates that independent corpus consultation is possible but cannot simply be assumed. Hua, Lu and Guo (2024) found that students experienced difficulties with query syntax, interpretation of corpus results and technical access, demonstrating the importance of corpus-literacy training.

For Philology students at Fergana State University, corpus literacy can therefore become a bridge between **digital pedagogy, English language development, learner autonomy and academic research competence**.

COCA, BNC, AntConc, LancsBox and SKELL can be used not merely as additional digital resources but as instruments for independent investigation. Through carefully scaffolded instruction, students can learn to formulate questions, search linguistic data, identify patterns, evaluate evidence, apply findings and reflect on their learning.

The central pedagogical principle can be summarised as follows:

Students should not only learn English; they should learn how to investigate English.

When students can independently investigate why a word is used, which collocations are conventional, how grammatical patterns operate and which expression is appropriate in a particular register, they develop a form of digital academic competence that extends beyond a single English course.

Therefore, integrating corpus literacy into Philology programmes at Fergana State University may provide a practical pathway toward more evidence-based, student-centred and sustainable independent English language learning.

Foydalanilgan adabiyotlar

[1] Casal, J. E., & Kessler, M. (2024). New theoretical and practical horizons in Data-Driven Learning: Introduction to the special issue. TESOL Quarterly, 58(3), 1027–1045. https://doi.org/10.1002/tesq.3331.

[2] Farr, F., & Leńko-Szymańska, A. (2024). Corpora in English language teacher education: Research, integration, and resources. TESOL Quarterly, 58(3), 1181–1192. https://doi.org/10.1002/tesq.3281.

[3] Hua, Y., Lu, X., & Guo, Q. (2024). Independent corpus consultation for collocation use in academic writing by L2 graduate students. System, 127, 103515. https://doi.org/10.1016/j.system.2024.103515.

[4] Li, D., Noordin, N., Ismail, L., & Cao, D. (2025). A systematic review of corpus-based instruction in EFL classroom. Heliyon, 11(2), e42016. https://doi.org/10.1016/j.heliyon.2025.e42016.

[5] Emir, G., & Saral, N. Ç. (2025). Exploring emerging issues in corpus applications in ELT: A systematic review. Participatory Educational Research, 12(2). https://doi.org/10.17275/per.25.16.12.2.

[6] Mukherjee, J. (2006). Corpus linguistics and language teacher education. In E. Braun, K. Kohn, & J. Mukherjee (Eds.), Corpus technology and language pedagogy: New resources, new tools, new methods. Peter Lang.

[7] O’Keeffe, A., McCarthy, M., & Carter, R. (2007). From corpus to classroom: Language use and language teaching. Cambridge University Press.

[8] Johns, T. (1991). From printout to handout: Grammar and vocabulary teaching in the context of Data-Driven Learning. In T. Johns & P. King (Eds.), Classroom concordancing. ELR Journal, 4, 27–45.

[9] Anthony, L. (2023). AntConc (Version 4.2.4) [Computer software]. Waseda University.