



Образовательный потенциал цифрового ресурса на основе корпуса студенческих текстов

Е. В. БОРЗОВА, И. А. КОТЮРОВА

АННОТАЦИЯ

Введение. Образование с использованием больших данных становится заметной особенностью в российских университетах, однако некоторые аспекты обработки и анализа больших данных требуют особых усилий при организации процессов обучения и учения, особенно по гуманитарным направлениям подготовки специалистов. *Целью данной статьи* было исследование образовательного потенциала нового цифрового ресурса, разработанного на основе корпуса ученических иноязычных текстов.

Материалы и методы. Основным объектом теоретико-методического анализа выступил лингвистический корпус студенческих текстов на иностранном языке, рассматриваемый как инновационный электронный образовательный ресурс. Теоретический анализ психолого-педагогической и методической литературы и обобщение практического опыта на материале Петрозаводского аннотированного корпуса текстов (ПАКТ) стали ключевыми методами исследования.

Результаты. Авторы определили методологическую базу для обоснования эффективности новой модели электронного образовательного ресурса: системно-деятельностный, личностно-ориентированный подходы, теория двойной петли обучения и принципы хьютагоики. Дизайн, содержание и вовлечение студентов в соответствующую деятельность опираются на положения о двойной петле учения, спиральной организации обучения и глубокой переработке усваиваемого материала. В 2021 году преподаватели кафедры немецкого и французского языков Петрозаводского государственного университета приступили к разработке особого цифрового ресурса – Петрозаводского аннотированного корпуса текстов (ПАКТ), который регулярно применяется в образовательном процессе для организации деятельности студентов в разных целях. Практика его использования доказывает эффективность при условии разработки качественного методического сопровождения, предназначенного для разных пользователей.

Заключение. Аналогичные цифровые ресурсы, разработанные на основе студенческих текстов или других продуктов учебной деятельности за период университетского образования и подкрепленные дидактическим инструментарием, можно разработать для создания среды, насыщенной большими данными, и по другим специальностям.

КЛЮЧЕВЫЕ СЛОВА

образование на основе больших данных, цифровой ресурс, корпус ученических иноязычных текстов, образовательный потенциал, методическое сопровождение

Для цитирования: Борзова Е. В., Котюрова И. А. Образовательный потенциал цифрового ресурса на основе корпуса студенческих текстов // Перспективы науки и образования. 2026. № 5. С. 252–265. <https://doi.org/10.32744/pse.2026.5.16>

Поступила: 07.04.2025 | Одобрена: 15.07.2026 | Опубликована: 31.10.2026



Лицензия «С указанием авторства — С сохранением условий». Позволяет перерабатывать, исправлять и развивать произведения при условии указания авторства и лицензирования производных работ на аналогичных условиях.



Educational Affordances of Learner Corpora Based Digital Resource

E. V. BORZOVA, I. A. KOTYUROVA

ABSTRACT

Introduction. Big Data Driven education is becoming a noticeable characteristic of modern university education in Russia, especially in the fields of engineering and medicine. *The aim of the current study* was to explore the teaching and learning affordances of a new digital resource developed on the basis of students' foreign language corpus.

Materials and methods. The primary object of theoretical and methodological analysis was the learner foreign language corpora used as a content base for a digital resource in the context of pre-service teacher education. The theoretical analysis of psychological-pedagogical and methodological literature, along with the practical approbation based on the Petrozavodsk Annotated Corpus of Texts (PACT), became the key research methods.

Results. The authors identified several learning theories that underlie the development of the Learner Corpora Based Digital Resource (LCBDR) and the guidelines for the student involvement into effective activities. The design and content of the LCBDR rely on the ideas of double-loop learning, spiral approach and recycling. The LCDDR named PACT was created 2021 by the German and French Languages Department of the Petrozavodsk State University and has been used on a regular basis for a variety of educational activities. The practical implementation shows that this digital resource proves to be an effective instrument for teaching and learning in case it is provided with a well thought out activity support guidelines for all those involved.

Conclusion. Similar digital resources based on students' generated texts or other products, collected throughout their university studies and backed up with appropriate didactic tools, can be developed for creating big data learning environments in different majors.

KEYWORDS

big data driven education, a new digital resource, students' foreign language corpora, the teaching and learning affordances, activity support guidelines

For Citation: Borzova, E. V., & Kotyurova, I. A. (2026). Educational Affordances of Learner Corpora Based Digital Resource. *Perspektivy nauki i obrazovaniya = Perspectives of Science and Education*, (5), 252–265. <https://doi.org/10.32744/pse.2026.5.16>

Received: 07 April 2026 | Approved: 15 July 2026 | Published: 31 October 2026



This is an open access article distributed under a Creative Commons Attribution-ShareAlike International License (CC-BY-SA 4.0) that allows others to share the work with an acknowledgement of the work's authorship and initial publication in this journal

INTRODUCTION

Education worldwide is currently undergoing digital transformation. UNESCO, UNICEF, ITU, the Global Partnership for Education as well as public and private partners of the Digital Transformation Collaborative, have developed a common framework – Six pillars for the digital transformation of education [1] – to guide this process. In line with the global educational initiatives, highlighted in the journal “Innovating Pedagogy” [2; 3], modern university education should focus on the students’ personal development, underlying the ability to learn continuously. Consequently, university foreign language education faces the challenge related to facilitating the students’ awareness of the personal significance of learning under the conditions of digital invasion. However, traditional electronic educational resources are often linear and text-heavy, offering little room for students to move beyond uniform, pre-set content and become genuine agents of their own learning. The UNESCO Institute for Information Technologies in Education (IITE) is actively advancing the concept of Smart Education [4], which leverages data and technology to personalize learning and improve outcomes. The use of linguistic corpora, including learner corpora, is an integral part of this approach.

Metalinguistics studies appear to provide the most systematic grounds for the effectiveness of the corpus based learning. A. Boulton and T. Cobb [5] confirmed that Data-Driven Learning (DDL) guarantees a sustainable positive effect on the language competence development, especially in terms of vocabulary and collocation acquisition. Some recent studies proved that DDL can be beneficial for the development of the student research competence in the field of linguistic theory learning, such as lexicology, for example [6], as well as for acquiring writing skills [7; 8].

Learner Corpora occupy a special place in corpus linguistics. Learner Corpora are defined as a collection of texts produced by foreign language learners themselves. It is evident that in DDL little emphasis is placed on Learner Corpora. An overview of the linguistic corpora types (2021) used in DDL scenarios reveals that only 4% of the conducted studies (21 out of 489) deal with learner corpora [9]. In the past, learner corpora were mostly used to identify typical errors and cases of mother tongue interference [10]. However, in recent years, there has been a shift from the deficit model to the resource model, when Learner Corpora are considered a source of information of what learners can do in the foreign language [11].

In the review of the learner corpora application for educational purposes conducted in 2024, S. Granger, head of the Learner Corpus Association, concludes that this application usually boils down to two major categories: a) comparison of the data obtained from the students’ texts and those from the native speakers’ or/and b) error correction [12]. Meanwhile, there has been practically no study where learner corpora serve as a media and personally generated content for recycling one’s own learning experiences.

The problem issue of the current study is related to the contradiction between the proven didactic potential of the learner corpora and lack of the grounds for designing the Learner Corpora Based Digital Resource (LCBDR) aimed at the systematic implementation of the learner corpora didactic affordances in university foreign language education. The goal of the article is to identify and describe the teaching and learning affordances of a new digital resource developed on the basis of students’ foreign language corpora, keeping in mind certain theoretical assumptions, based on a few learning theories. The authors made an attempt to reveal appropriate grounds that underlie the development of the LCBDR intended for pre-service training of university foreign language teachers and then design the target resource implemented in the Petrozavodsk State University (Russia).

The current study explores the teaching and learning affordances of a new digital resource developed on the basis of students’ foreign language corpora. The texts written by students – future foreign language teachers – were used as a basis for the learners’ activities to achieve integrated goals: simultaneously enhance their foreign language, key, professional and research competences.

The authors identified several learning theories that underlie the development of the Learner Corpora Based Digital Resource in the literature review.

LITERATURE REVIEW

The current research is based on the theory of double-loop learning, proposed by C. Argyris and D. Schön [13]. Unlike single-loop learning, which relies on rules and error correction, double-loop learning involves attempts at looking into the previously acquired experiences from a different perspective. It entails a shift in the mental models and a change in the learner action strategies. Therefore, dealing with the primary experiences, the learner tries to get a better understanding of his/her activities and modify them accordingly to achieve quality outcomes, avoiding hasty choices, reproduction of inappropriate patterns, and erroneous tactics [13].

The concept of recycling in learning is closely connected with the theory of double-loop learning. Recycling in this case is defined as a process of repeated reactivation and application of the primary experiences under new circumstances and in new combinations to create a new product [14]. In foreign language teaching, recycling implies intensive practices in reusing previously learned language units in new situations which leads to a higher level of acquisition due to the development of speech acts flexibility [15]. Usually, methodologists focus on vocabulary recycling, highlighting a positive effect of frequent revision on sustainable vocabulary sub-skills [16; 17]. Meanwhile, G. Conti emphasizes grammar recycling extended for a longer teaching period, grounding this assumption on cognitive psychology, stating that system language acquisition and storing the material in memory require its multiple and variable usage which is a rare case in the present foreign language course books [18].

In our study, the choice of the experience recycling concept as a basic strategy for the LCBDR model is consistent with the assumption that quality acquisition depends not only on the number of repetitions and allocated time, but also on the cognitive efforts of the learner as well as on the ability to accumulate experience components to complete new tasks [19]. Prior experience analysis activates the basic mechanisms of recycling – thinking (its strategies and varied operations) and memory, which both promote accuracy in foreign language application and quality speech acts production in foreign language teaching and learning. Therefore, well-grounded and properly implemented recycling leads to the transformation and systematization of primary experiences, resulting in gaining secondary experiences as a higher level learning outcome. It is evident that the implementation of the recycling strategy requires specially designed aids, teaching technologies, and tasks.

In foreign language teaching and learning theories in Russia, there are scarce papers exploring the double-loop learning framework. This idea was implemented in the context of designing individual learning routes for university teacher in-service education [20] and teaching methods to foreign language university students in pre-service teacher training of [15].

Alongside with the double loop research, we find studies devoted to heutagogy – an approach of self-determined learning based on taking responsibility for the process and outcomes of education [21; 22]. The ideas of self-directed learning, gaining personal experiences as a basis for new life skills, and flexibility in decision making underlie this theory. L. M. Blaschke and S. Hase stress the key characteristics of heutagogy: learner agency (motivation and responsibility), capability (using one`s competencies in diverse situations), self-reflection and metacognition (awareness of how one is learning), double-loop learning (psychological and behavioral engagement) and nonlinear learning and teaching [23]. The implementation of the principles mentioned above suggests learners` active exploration, connection, collaboration, sharing, creation and reflection which are essential heutagogic design elements, according to the authors. Heutagogic principles are clearly intersected with the ideas of learner-centered approach, that is why heutagogy can be regarded as its branch.

We can witness a strong trend in modern education to place learners in a dominant position in education through creating engaging environments for their personally meaningful activities. In our study, it is claimed that such environments can be created by means of learner corpora and that the LCBDR model is one of the effective means to implement double-loop learning and recycling ideas in pre-service foreign language teacher education. To this end, it is possible to apply a tripartite framework suggested by A. Leńko-Szymańska which consists of corpora literacy, corpora-oriented language pedagogy and corpora-oriented reflexive practice [24]. This framework is empirically confirmed by Q. Ma et al. who proved that corpora literacy appears to be a key condition for the effective realization of corpus-based language pedagogy which contributes to

quality teaching outcomes [25]. The interlanguage perspective is also important for our study. Relying on bidirectional learner corpora of the Chinese and Japanese languages, K. Mochizuki et al. reveal the affordances of cross-referentiality of the two languages in the acquisition of some grammar categories (those of existence and possession structures) [26; 27]. It follows that the new LCBDR model should promote the development of the learners' corpora literacy in creating the resource for their further learning and reflexive practices.

Finally, it is necessary to consider the role of corrective feedback in corpus application. Y. Sataké, based on COCA (Corpus of Contemporary American English), concluded that the effectiveness of the corrective feedback is related to the source of the feedback itself. The teacher's feedback works exceptionally well to correct skipping and matching errors, while the peer feedback as well as self-assessment shows different results [28]. C. Sarré et al. conducted their studies in the same vein, comparing the impact of different assignments based on annotated learner corpora. They came to the conclusion that articles and verbal agreement sub-skills are developed better through the comparison of learner generated texts with those of native speakers (in their case The British National Corpus), though for verb tenses the traditional corrective feedback turned out more appropriate [29].

It is obvious that learner corpora offer valuable content for varied student activities which promote language accuracy as well as flexible language usage provided, they engage learners in cognitively challenging and appropriate tasks.

MATERIALS AND METHODS

The study presents a theoretical and methodological analysis of the ground assumptions identifying the affordances of the learner foreign language corpora used as a content base for a digital resource in the context of pre-service teacher education. The key study methods are a theoretical analysis, designing a digital resource model for educational purposes, and practical approbation method.

To achieve the aims of the study, a theoretical analysis of the related literature on psychology, pedagogics, methods of teaching and learning foreign languages and corpus linguistics was conducted. The methodological ground approaches of the study are the system activity-based approach and the learner-centered approach as well as the theories of double-loop learning and heutagogic principles.

The hypothesis of the current study is that the affordances of the Learner Corpora Based Digital Resource (LCBDR) consist in the achievement of a variety of integrated educational goals which go beyond sole error identification and error correction, provided the digital resource engages students in double-loop learning and primary experiences recycling in line with the heutagogic principles. To this end, the target LCBDR needs to be supplemented with a special teaching and learning support section, containing detailed instructions and diverse tasks for appropriate student activities.

The designed model of the Learner Corpora Based Digital Resource (LCBDR) called PACT (Petrozavodsk Annotated Corpus of Texts) has been implemented in the educational process since 2021. The outcomes of the approbation were thoroughly analyzed and briefly presented in the article.

RESULTS

Based on the theoretical analysis of the related literature, it is obvious that the key grounds of the LCBDR model under study are in line with the modern trends of university education, implying a shift from a deficit model to a resourceful one, from a single correction loop to a double learning loop, from self-centered, linear application to integrated application in the pre-service foreign language teacher training by means of reflexive practices development. The intersection of these characteristics reveal the specific features of the educational potential of the LCBDR as a new digital resource, compiled of the products created by students and intended for a variety of educational goals. In fact, this type of the LCBDR can be adjusted for any university course, given the specificity of the academic subject is taken into consideration.

The concept of the learner corpora digital resource

In this study, the learner corpora are defined as a digital collection of authorized written papers in a foreign language compiled at the university in accordance with certain requirements and supplemented with service functions that go beyond the traditional data search, providing visualization, task and test generation, error annotation, etc., in fact, multifaceted teaching and learning support.

The target learner corpora possess two types of features.

Content features include:

- Accumulation of diverse data, annotated for various types of analysis and activities.
- The context that is as close to the student as possible (texts produced by the student and groupmates).
- Teaching and learning support services which facilitate the application of the LCBDR for educational goals.
- Long term fixation of the learning dynamics which continues throughout the entire period of university studies.

Technical features include:

- Downloading texts supplemented with attribute information.
- Manual error marking and assessment.
- Automatic task and test generation.
- Search by words and types of errors.
- Statistics accumulation in various sections.

It is due to the fusion of these features that the target LCBDR appears to be not just a data collection, but a basis for designing a new digital educational resource. Its psychological and didactic specificity is that the students' activities are based on their own written texts which are products of their previous learning (their accumulated primary experiences). At the primary stage, learning unfolds according to the scheme "educational environment and related prerequisites – student activities – outcomes". The primary stage is aimed at the development of basic foreign language sub-skills and skills which are explicit in the texts produced by the students. These texts reflect the learners' primary experiences with all possible errors and deviations from standards. When in some time students come back to their own texts to complete new assignments on their basis, they get an opportunity to recycle what they have already learned, that is to rethink and transform their primary experiences. Unlike the majority of digital resources which focus on learners' primary experiences, the target LCBDR places emphasis on recycling, thus providing student quality outcomes (their secondary experiences, recycled and enhanced).

The LCBDR is characterized by a considerable role of the teacher who actively directs the production and collection of the learner corpora. However, gradually students take on responsibility and become more autonomous while working with the resource. These features overlap with the principles of heutagogy. It is also obvious that due to the learners' engagement in diverse activities with the corpora, the idea of recycling the learners' primary experiences is consistently implemented which promotes double-loop learning, thus facilitating a noticeable improvement of the outcomes and creating their secondary experiences.

Unlike many traditional digital resources, the target LCBDR implements the idea of multiple student activities with the same materials for varied goals. The learning process presents a spiral. The concept of the spiral learning model was first proposed by J. Brunner according to which every subject or skill area is revisited at intervals, at a more sophisticated level each time. Foreign language acquisition based on this concept definitely enhances the quality of the skills and competences that students develop, promoting "a spiral of increasing ability in using language in order to get something done" [30].

The ground idea of the target LCBDR is that learners keep working with the components of their own experiences, dealing with them from different perspectives: searching for data, analyzing information, systematizing it, introducing changes or corrections, etc. Therefore, due to this approach, it is possible to overcome fragmentation in learning, ensuring continuity, connections and flexibility in the development of speech mechanisms. The learners' primary experiences gained at the point of presentation and first stage practices, improve because of the emergence of new associations as well as strengthening patterns and links, facilitating vocabulary and grammar expansion which are intentionally recycled in the subsequent stages. Learners have an opportunity to repeatedly analyze and use the language in new situations for the discussions of different topics and fulfillment of diverse language, communicative and research tasks. We can expect that involving learners in systematic activities with the target LCBDR, compiled of their own texts and supplemented with a variety of related assignments, will promote clarity and better awareness of the language. These gains will be revealed in the accurate language usage, on the one hand, and flexibility and adjustment of speech mechanisms, on the other.

In this respect, we can cite D. Eagleman, an American neuroscientist, who states that the XXI century education must shift from rote memorization to learner's cognitive flexibility, developed in creative student activities, encouraging research culture through critical thinking and problem-based learning. He claims that learning to "dig deeper" is based on the analysis of the first responses which can be erroneous. Errors are inevitable, but they can be used for reflection, facilitating more profound knowledge and more flexible skills in an environment that inspires learner's creativity and self-fulfillment [31].

Learner corpora can obviously contribute to the creation of such an environment and stimulation of such activities because students interact with their own experiences, not only positive ones at that. Delayed return to their personal outcomes, aimed at their thorough analysis and new applications, teaches learners to identify some repeated faulty choices and discover possible reasons for errors, shaping a view from another perspective. It is also evident that the recycled subject-matter is meaningful for learners which is highly likely to place them into the agentic position in their interaction with the content of the target LCBDR.

Summing up, we conclude that learners' foreign language corpora can be transformed into a new model of a digital resource provided it is supplemented with an appropriate teaching and learning support oriented at integrated goals and provided with consistent tasks and instructions which will fulfill its educational affordances to the full.

Practical outcomes of the development and application of the target LCBDR

The authors' practical implementation of the theoretical grounds described above confirms the preliminary assumptions. PACT (Petrozavodsk Annotated Corpus of Texts) [32] falls into the category of innovative digital resources, having changed from an ordinary students' foreign language corpora into the resource aimed at a variety of educational purposes.

Its content characteristics, which allow to implement a number of diverse functions, consist in containing big data, annotated and ready for analysis as well as diverse research work (comparative, interferential, interdisciplinary, lingua-didactic, psycholinguistic, etc.); accumulating written papers throughout the entire period of university studies; the closest possible context; teaching and learning support supplement which directs students' activities for diverse goals.

The technical characteristics of PACT promote its consistency with the polyparadigmality principle that the perspective students' and teachers' activities are based on: the possibility of users' authorization either as a student or a teacher in order to complete diverse activities (downloading, attributive marking, manual error marking and assessment, automatic task and test generation related to the identified errors, search for relative sentences by words and types of mistakes, collecting students' texts and providing their secondary recycling, statistics and results accumulation in various sections).

The characteristics mentioned above reveal the specificity of PACT as a LCBDR sample.

The PACT content contains data consisting of students' written texts downloaded since the first year of their university studies up to their graduation. "Personally produced papers and those written by their group-mates create the closest context which is understandable, personally meaningful and clear to every student both from the language and culture related aspects", whereas "material adjustment to the cultural background as well as to the learners' interests and needs encourages their engagement into learning" [33].

In many cases this type of content turns out to be exceptionally appropriate for data driven learning or corpora-based approach as the sentences displayed in the corpora search results seem to be adapted for particular learners. Consequently, in the process of the reception, the students' attention is not distracted by the authenticity of complex semantic or unfamiliar syntactic units, being able to focus on currently significant language phenomena. However, it should be noted that in the majority of learning situations, students' texts must not replace authentic German language materials. The learner corpora are regarded as a supplementary resource, another additional tool that makes sense to use in compliance with the set objectives, as is the case of a new topic introduction or revision/recapping the previous language/topic or error correction or, finally, for doing project or research work.

The uniqueness of the learner corpora as a learning resource also consists in the fact that collecting learning products during the entire period of university studies provides a chance to objectively track personal dynamics and, based on it, design a personal learning trajectory. As was already stated, goal-oriented activities with one's own texts completed from a different perspective increase their personal importance for students and potentially place them in the agentic position of their learning. This position implies striving for better outcomes of foreign language acquisition, planning adequate steps for that, reflecting on one's prior experiences from today's perspective and simultaneously improving self-control and self-assessment mechanisms.

Learning a foreign language based on learner corpora is also noted for the strategy "learning from mistakes, both one's own and other's". Above, we pointed out the importance of erroneous experience in learning, highlighted by neuroscience (D. Eagleman). However, there are still controversies among researchers and practicing teachers in relation to student's errors. In Russia, K.G. Chiknawerova [34] conducted a profound study of different approaches to errors made by foreign language learners. In the diachronic context and at the present stage, the researcher describes radically opposite theories, ranging from absolute rejection of correction measures to attributing all possible benefits to them. In this regard, the researcher concludes that "anyway, even nativist theories proponents recognize the appropriateness of error correction in certain cases. Despite some doubts regarding traditional foreign language teaching methods, at present these methods are being reviewed and their efficiency is being justified in modern educational contexts [34].

In PACT, error annotation in students' texts is not an end in itself. Errors are inevitable in foreign language learning. It is not only and not so much about correcting mistakes, but rather about content selection for thorough practicing and about choosing appropriate methods aimed at their radical eradication and subsequent prevention. Students' errors can be applied to encourage reflection, to design individualized visual aids and automatic prompts, to generate tasks and tests, etc. The point of activities which are based on the identified errors is recycling the meaningful content to achieve integrated quality outcomes.

Thus, for example, the task "annotate errors in the corpora" facilitates the simultaneous achievement of a few objectives: enhancing language sub-skills as well as developing the professional competence of the future teacher whose activities suggest identifying errors, defining their types and degree of crudeness, and, which is more important, finding acceptable ways of explanation and effective practices. Undoubtedly, these activities are beneficial both for language learners and teachers as well because they allow students to improve their own language sub-skills and, in addition, develop supervising and corrective functions of the foreign language teacher. Different annotation assignments related to one's own texts as well as to other students' texts have proved to be useful in this respect if they are proposed on a regular basis and completed in different ways of student interaction (individually, in pairs or groups). This is only one example demonstrating the possibility of simultaneous development of different competences, whereas PACT contains a great

variety of tasks, facilitating interconnected enhancement of language, communicative, universal and professional competences on the basis of the collected learner corpora.

Another specific feature of the learner corpora based PACT is the service of task and test generation in view of the students' personal characteristics and outcomes. These tasks and tests are correlated with the particular learner's errors. Consideration of the charts which reveal the number of errors of varied types made by every student, provides information, concerning individual weaknesses in language acquisition. It allows to offer personally oriented tasks based on the sentences borrowed from of the students in the group with the similar error type. Respectively, the learner corpora shape the content basis for all language sub-skills enhancement (spelling, vocabulary, and grammar), promoting the fixation of the correct linguistic patterns in the long-term memory and also contributing to the improvement of the mechanisms of their usage in both oral and written communication.

The automatic task and test generation tools installed in PACT turned out to be both effective and easy to use increasing the degree of the students' agency as language learners and future foreign language teachers as well. PACT allows students to generate tasks and tests on their own, making reasonable choices of appropriate sentences from their own or group-mates' texts and omitting irrelevant ones. As a result, the learners' activities become personally oriented because they are aimed at the recycling of their own primary experiences to achieve better outcomes in their foreign language learning.

PACT services allow teachers to identify apparent trends in the learners' progress of each group and also keep them informed of particular error types of both the whole group as well as of every student. It enables teachers to design the learning curriculum in line with the statistics, taking into account real time issues and adjusting the learning activities to match the current state of things. From the point of view of the student, the prompt feedback from the resource enables him or her to keep up with how the learning process occurs. Learners have an opportunity to analyze both their achievements and drawbacks in the language acquisition, gaining positive as well as negative experiences for reflection. It is evident that the resource materials encourage the students' critical thinking and self-assessment skills, aimed at identifying problem issues of their learning. Learners discover real life boundaries and limitations of their knowledge or ignorance, high or low proficiency level, success or failures. This information helps design their subsequent activities with a view of relevant study skills development and awareness of the value of personal achievements.

The materials collected in PACT are structured and supplemented with search tools as well as other elements inherent in corpus managers, thus enabling users to implement the search-activity approach to university foreign language education. Using the resource, students gain experiences under the teacher's supervision who sets tasks and guides the learner's activities, providing all kinds of support if it is needed. These teacher's functions are especially important in the first year of the university studies because students are involved in working with a new resource and they need detailed instructions. Gradually, they get used to this type of activities and do tasks either in the classroom or independently.

A significant advantage of PACT is its open character. The number of the students' corpora collected in the resource keeps growing constantly which expands its educational potential and requires further improvement of both its content and technical affordances. The scope of content changes may include, for example, additions of different text genres or creation of sub-corpora sections that contain texts written by students from other universities or by Russian or other language speakers, as well as advancement of the annotation services of the resource. Another possibility to enhance the resource content functions is to introduce new criteria for statistical outcome processing. From the technical perspective, it is possible to expand its functions with the help of Artificial Intelligence, thus opening up a new space for mastering the lingua-didactic aspects of the resource.

Another notable peculiarity of PACT is that, with its expansion, it is possible to incorporate in it two apparent blocks: archival which contains texts produced by university graduates and operational which contains texts of those who are currently studying and working actively with the corpora. The archive block preserves all the texts from the previous years and is mostly used for linguistic and pedagogical research purposes. The operational block is intended first of all for the fulfillment of learning objectives by its developers who gradually write and download their texts, simultaneously completing diverse tasks on their basis.

The description of PACT reveals the importance of the instructional and learning support for students which facilitates the quality of their studies, on the one hand, and aids teachers to guide the educational process efficiently. Without the provision of this function, the resource would not be able to implement its educational affordances. It is evident that effective learning tools must always be compatible with new educational goals and innovative content, new methods, forms, and teaching techniques, being based on certain teaching principles" [35]. Therefore, it is of primary importance for researchers to identify such principles, grounding them on solid learning theories, and suggest reliable guidelines for arranging adequate teaching and learning activities based on learners' foreign language corpora. An essential part of the guidelines must be an expanding set of tasks of varied complexity and algorithms of the learners' activities which guarantee quality learning outcomes. This is the key purpose and prerequisite for designing effective learner corpora based resource in teaching and learning languages elsewhere.

DISCUSSION

The description of the PACT ground ideas and their implementation confirms the hypothesis that the affordances of the learner corpora based digital resource are more versatile and multifaceted than those of traditional linear digital resources. They also clearly distinguish this corpora type from most corpora managers. This type of the resource allows to shift the educational emphasis from the presentation of the "correct" sample acquisition to encouraging learner's active engagement in recycling personally meaningful content, presented in the learner's product/text (primary experiences). Thus, PACT implements the spiral idea of learning proposed by J. Bruner and in our case applied to university foreign language acquisition in the context of pre-service foreign language teacher education. As I. Bryxina and A. Sidorova [36] point out, the spiral model is quite rarely used for teaching foreign languages, though, according to the authors, it is essential for foreign language grammar learning. The activities based on the students' prior knowledge facilitate a deeper understanding of the language rules and sub-skills automation. We share the assumption made by I. Bryxina and A. Sidorova that the inherent characteristics of the spiral model determine its effectiveness [36]. A.V. Begunts and O.S. Solov'eva [37] explain such developmental effects of the spiral model through the learners' gradual, multi-layered understanding of the target content which is quite evident in case of PACT when each new tasks based on the resource encourage a cyclic return to the previously practiced materials. Students repeatedly recycle their whole previous experiences of foreign language acquisition, each time with a new focus. This strategy enables them to improve a wide variety of sub-skills and skills. As our approbation proved, the spiral model turned out to be a productive alternative to the linear model of foreign language learning.

A frequent question asked by teacher users of PACT is that of the negative impact of focusing learners on the texts which are far from being ideal. There is no denying the fact that the products the students work with in PACT inevitably contain language errors or other drawbacks. The approach underlying the resource based on error-driven learning strategy is consistent with the ideas expressed by L. Eskreis-Winkler and A. Fishbach [38] who experimentally proved that the majority of people are not able to take advantage of their mistakes. For many of us negative experiences are discouraging and demotivating. However, the authors state that "the information in failure is often high-quality information that can be used to predict success" [38], though learning by mistakes needs special practice. That is what PACT provides. Coming back to their own products again to complete new tasks gives impetus to reflection, significant for transforming as well as mastering the experiences gained before, because it suggests not only reproduction of the correct sample, but a profound review of the grounds for one's own actions and choices. Therefore, it strengthens the learner's agentic position in experiences recycling and deep learning. These characteristics of the student's activities are compatible with the latest trends of the XXI century education. The double-loop learning theory which has not been widely applied before comes to the forefront of grounding the design of the learner corpora based resource. PACT as a learner corpus based digital resource makes the traces of the student's previous actions vivid and available for thorough analysis, assessment and recycling.

The approbation of PACT resource in the Petrozavodsk State University proves that teachers get at their disposal a wide range of diverse didactic tools and services for encouraging and supervising learners' activities. Students get impetus to work systematically with the resource to achieve quality learning outcomes in university pre-service foreign language teacher education. The key directions of the teachers' activities are

- providing a set of varied assignments and activity algorithms, applied on a regular basis and distributed throughout the whole period of university studies;
- developing mechanisms of error correction and error prevention (self-control mechanisms);
- facilitating reflection and recycling of one's primary experiences (accumulating and refining one's experiences).

In the senior years of university studies, these language learning functions are supplemented by those aimed at the development of the research competences in line with the goals of theoretical disciplines (both linguistic and pedagogical). Overall, all the tasks and activity types also contribute to the simultaneous development of universal and professional competences.

The students, in their turn, get access to a variety of materials based on corpora and diverse data, relevant to the awareness of their own dynamics in language learning. Completing meaningful tasks and identifying problem areas in language acquisition on a regular basis also result in the development of coping strategies and tactics based on reflection. Students improve their skills of information search and analysis as well as their systematic application for diverse tasks and problem solution, both in individual and team work. It is obvious that these tasks promote mastering critical thinking and interactive skills as well. Finally, in addition to learning the language, the learner based corpora resource turns into a tool of completing feasible research tasks, offering varied assignments which suggest hypothesizing, generating argumentation and solutions as well as doing project work and participating in discussions. Working with the personally relevant texts over time, students definitely enhance their well-grounded self-evaluation skills.

However, we should not disregard certain limitations of PACT. Firstly, the application of the resource requires high level digital literacy of the university teacher. Secondly, it takes time to overcome possible psychological student resistance as not each of them is ready to return to his/her own errors in a formal learning situation. Nevertheless, in case the resource is backed up with an appropriate teacher and learner support guide, this resistance is likely to notably decrease, especially when the student becomes aware of the evident benefits of the regular resource usage, resulting in the growing language accuracy and speech fluency.

Another important debatable issue is that of the ratio of the archive and operational blocks in PACT. The archive block builds the basis for comparative research (such as, for example, identifying common errors in different study periods), but there are certain ethical rules of its application. Meanwhile, the operational block proved to be more dynamic and learning oriented.

Nowadays, the most burning issue is the role of Artificial Intelligence platforms in education. In this resource, AI can undertake the functions of automatic assessment, or prompt, or partial text generation, but then the risks of downloading AI generated texts may definitely increase. To avoid this problem, it is possible to permit the AI application only in its supporting function which can be detected by means of the following expertise and transparent recheck by the teacher and other students as well. Undoubtedly, it is crucial to foster learners' honesty and responsibility in terms of text production, which is rather a pedagogical issue, not a technical one.

Comparing PACT with foreign analogues, such as Cambridge Learner Corpus (URL: https://www.wfp.cupeng.org/elt/corpus/learner_corpus.htm) or Fehlerannotiertes Lernerkorpus (URL: <https://hu-berlin.de/falko>), shows that the majority of learners' corpora present research tools and are rarely in demand in everyday educational practices. The learner corpus based resource PACT is designed with an emphasis on integrated educational goals, distributed throughout a long period of university foreign language studies. The practical application of PACT confirms its enormous educational affordances.

CONCLUSION

The learner corpora possess a unique blend of content (a collection of foreign language texts produced by students characterized by a close context and learning environment, encouraging personally meaningful activities and monitoring real life learning dynamics both retrospectively and currently, providing statistical summaries, etc) and technical features (annotation services, information storage and search, task and test generation on demand, visualization services and statistics generation, etc).. This blend sets it apart from other digital or standard corpora managers.

The key psycho-pedagogical grounds for PACT are the concept of the double-loop learning and recycling one`s own primary experiences, thus providing the spiral learning model. The corpora fixes the learner`s primary experiences which then is reflected and transformed through further student activities, based on the collected texts with a shift of task focus. As a result, students develop a higher level sub-skills, skills and competences which accumulate their secondary experiences of higher quality. The holistic resource model is based on learner foreign language corpora, systematically implementing the ideas of personal engagement, spiral model of learning trajectory, and recycling one`s experiences to achieve educational quality outcomes.

The development of the LCBDR is based on three groups of principles; *organizational and technical* (copyright, data protection, diverse services and functions); *didactic* (learner orientation, interaction, variety of appropriate activities), and *specific* (openness, multiple experience inventory/recycling, multifunctionality, reflection).

Further study of the LCBDR can include experimental approbation of the model effectiveness for identifying cross-referentiality of different language pairs; the development of the automatic assessment system partly based on AI; identifying ways to preserve the "authenticity" of the learner created texts in the AI era; comparison of the teaching and learning outcomes achieved by means of traditional resources and the LCBDR; exploring affordances of teaching other academic subjects with the help of the LCBDR.

To implement the educational affordances of the LCBDR to the full, it is crucial to develop adequate teaching and learning guidelines and also provide relevant technical services. This type of resource may potentially change the whole concept of teaching and learning. The learner corpora in this case may turn into a stimulating learning environment, being based on personally created texts, easy to understand and personally relevant, when errors are not regarded as a disadvantage, but become the impetus for further growth and excellence.

REFERENCES

1. Six pillars for the digital transformation of education: a common framework. Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000391299> (accessed 19 May 2026)
2. Innovating Pedagogy 2023. The Open University Innovating Pedagogy Report Series. Available at: <https://iet.open.ac.uk/files/innovating-pedagogy-2023.pdf> (accessed 19 May 2026)
3. Innovating Pedagogy 2024. The Open University Innovating Pedagogy Report Series. Available at: <https://iet.open.ac.uk/files/innovating-pedagogy-2024.pdf> (accessed 19 May 2026)
4. UNESCO. Theme: Smart Education. Available at: <https://iite.unesco.org/theme/smart-education/> (accessed 19 May 2026)
5. Boulton A., Cobb T. Corpus Use in Language Learning: A Meta-analysis. *Language Learning*, 2017, vol. 67, no. 2, pp. 348–393. DOI: 10.1111/lang.12224
6. Kachurin I.D. Application of Corpus Technologies in Teaching Linguistic Theory (on the example of German word formation). *Scientific Opinion*, 2024, vol. 9, pp. 150-164. DOI: 10.25807/22224378_2024_9_150 (In Rus.)
7. Anthony L. Introducing corpora and corpus tools into the technical writing classroom through data-driven learning (DDL). *Discipline-specific writing*, 2016, pp.176–194. DOI: 10.4324/9781315519012-18
8. Şahin Kızıl A. Data-driven learning: English as a foreign language writing and complexity, accuracy and fluency measures. *Journal of Computer Assisted Learning*, 2023, vol. 39, no. 4, pp. 1382–1395. DOI: 10.1111/jcal.12807
9. Boulton, A., Vyatkina, N. Thirty years of data-driven learning: Taking stock and charting new directions over

- time. *Language Learning & Technology*, 2021, vol. 25, No.3, pp. 66–89. DOI: 10.64152/10125/73450
10. Callies M. Current perspectives on Learner Corpus Research. *AAA: Arbeiten aus Anglistik und Amerikanistik*, 2023, vol. 48, no. 1, pp. 37–52. DOI: 10.24053/AAA-2023-0002
11. Staples S., Dang A., Wang H. Learner corpora in corpus-informed instruction: Moving toward an asset- and genre-based model. *TESOL Quarterly*, 2024, vol. 58, no. 3, pp. 1166–1180. DOI: 10.1002/tesq.3293
12. Götz S., Granger S. Learner corpus research for pedagogical purposes: An overview and some research perspectives. *International Journal of Learner Corpus Research*, 2024, vol. 10, no. 1, pp. 1–38. DOI: 10.1075/ijlcr.00039.got
13. Argyris C., Schön D. *Theory in Practice: Increasing Professional Effectiveness*. San Francisco: Jossey-Bass; 1974, 224 p. Available at: <https://archive.org/details/theoryinpractice0000argy/page/n5/mode/2up> (accessed 19 May 2026)
14. Helgesen, M. One more time, with feeling—Bringing task recycling to the classroom. *KoTESOL English Connection*, 2003, vol. 7, no. 3, pp. 1–8.
15. Borzova, E. Recycling in teaching English as a major. *Foreign Language Education in a Multicultural Environment: Proceedings and Reports of the XXIV Scientific-Practical Conference of the National Association of Teachers of English (NATE2018)*. Samara, 2018, pp. 76–78.
16. Zahar R., Cobb T., Spada N. Acquiring Vocabulary through Reading: Effects of Frequency and Contextual Richness. *Canadian Modern Language Review*, 2001, vol. 57, no. 4, pp. 541–572. DOI: 10.3138/cmlr.57.4.541
17. Muhammad A., Zaheer H., Azhar B., Muhammad I. Recycling of vocabulary in English language teaching: from theory to practice. *Epistemology*, 2020, vol. 7, no. 8, pp. 88–102.
18. Conti G. It is not just about 'how often', but 'how' you recycle – five learning principles that make it or break it in L2 grammar instruction, 2018. Available at: <https://gianfrancoconti.com/2018/01/04/it-is-not-about-how-much-but-how-you-recycle-five-learning-principles-that-make-it-or-break-it-in-l2-grammar-instruction/> (accessed 19 May 2026).
19. Glebova M.V. Neurobiological aspects of increasing cognitive effectiveness of learning in a modern general education school. *Modern high technologies*, 2023, vol. 12, pp. 276–283. DOI: 10.17513/snt.39894
20. Shemanaeva M.A. Individual learning path construction principles in the framework of further language education programs. *The Herald of South-Ural state Humanities-Pedagogical University*, 2015, vol. 6, pp. 107–113. (In Rus.)
21. Hase S., Kenyon C. From Andragogy to Heutagogy. UltiBASE In-Site. 2001. Available at: <http://pandora.nla.gov.au/nph-wb/20001220130000/http://ultibase.rmit.edu.au/Articles/dec00/hase2.htm> (accessed 19 May 2026).
22. Ignatovich E. Heutagogy as a foreign concept of self-determined learning. *Lifelong education the XXI century*, 2013, vol. 3, no. 3, pp. 108–116. DOI:10.15393/j5.art.2013.2151 (In Rus.)
23. Blaschke L.M., Hase S. Heutagogy: A Holistic Framework for Creating Twenty-First-Century Self-determined Learners. *The Future of Ubiquitous Learning. Lecture Notes in Educational Technology*. Berlin, Heidelberg: Springer; 2016, pp. 25–40. DOI: 10.1007/978-3-662-47724-3
24. Leńko-Szymańska, A. Training teachers in corpus-based language pedagogy: A tripartite framework. *Corpus Linguistics and Translation Studies*, 2022, pp. 509–524. DOI: 10.4324/9781003002901-39
25. Ma Q., Yan J., Crosthwaite P., Lee C. K. J. TPACK for corpus technology: TESOL student teachers' growth and self-efficacy in independent language learning and teaching. *ReCALL*, 2026, vol. 38, no. 1, pp.1–17. DOI: 10.1017/S0958344025100347
26. Mochizuki K. Shirai Y. The Acquisition of Aspect in Chinese Based on Learners' L1 Typology: An Analysis Based on the TUFs Co-referential Learner Corpora of Chinese and Japanese. *Learner Corpora: Construction and Explorations in Chinese and Related Languages*, 2023, pp. 229–256. DOI: 10.1007/978-981-19-5731-4_11
27. Fukuda S., Zhang Z., Mochizuki K. Acquisition of existence and possession structures in Japanese-Chinese cross-referentiality of two-way learner corpora: Beginner and advanced level-wise production trends and analysis. *Learner Corpus Studies in Asia and the World*, 2024, vol. 6, pp. 81–102. DOI: 10.24546/0100487716
28. Satake Y. The effects of teacher, peer and self-feedback on error correction with corpus use. *Applied Corpus Linguistics*, 2024, vol. 4, no. 3. DOI: 10.1016/j.acorp.2024.100114
29. Sarré C., Brudermann C., Grosbois M. Using learner corpus data for grammatical accuracy development in written productions: The role of corrective feedback. *International Journal of Learner Corpus Research*, 2024, vol. 10, no. 1, pp. 107–145. DOI: 10.1075/ijlcr.00042.sar
30. Edge J. Cross-cultural paradoxes in a profession of values. *TESOL Quarterly*, 1996, vol. 30, no. 1, pp. 9–30. DOI: 10.2307/3587605
31. Eagleman D. *Livewired: The Inside Story of the Ever-Changing Brain*, 2021, 336 p. (In Rus.)
32. Kotiurova I.A., Shchegoleva L.V. Learner Corpus in German as a Data Source for Education and Science. *Educational Studies Moscow*, 2022, vol. 4, pp. 322–349. DOI: 10.17323/1814-9545-2022-4-322-349
33. World Education Trends in the Russian Context 2025: A Study by Ultimate Education and HSE University. Available at: https://ioe.hse.ru/innovations/trends2025_an (accessed 19 May 2026).
34. Chiknawerova K.G. Interpretation and correction of errors in the speech of foreign language learners: A review of foreign studies. *Kazan Pedagogical Journal*, 2018, vol. 3, no. 128, pp. 129–136.
35. Shamov A.N. Foreign language teacher and his tools methodological activities. *Tambov University Review*.

- Series: Humanities*, 2025, vol. 30, no. 4, pp. 857–873. DOI: 10.20310/1810-0201-2025-30-4-857-873
36. Bryxina I.E., Sidorova A.A. Spiral curriculum in teaching French grammar for future linguists. *Tambov University Review. Series: Humanities*, 2024, vol. 29, no. 2, pp. 438-451. (In Russ.) DOI: 10.20310/1810-0201-2024-29-2-438-451
37. Begunts A.V., Solov'eva O.S. On the application of the didactic spiral in the construction of educational programs. *Lomonosov Pedagogical Education Journal*, 2021, no. 4, pp. 15-36. (In Russ.) DOI: 10.55959/MSU2073-2635-20-2021-4-15-36
38. Eskreis-Winkler, L. and Fishbach, A. You Think Failure Is Hard? So Is Learning From It. *Perspectives on Psychological Science*, 2022, vol. 17, no. 6 DOI: 10.1177/17456916211059817

Авторы

Борзова Елена Васильевна

(Россия, Петрозаводск)
Профессор, доктор педагогических наук,
профессор кафедры английского языка
Петрозаводский государственный университет
E-mail: borzovaelena40@gmail.com
ORCID: 0000-0002-5006-1480
Scopus Author ID: 57210163835

Котюрова Ирина Аврамовна

(Россия, Петрозаводск)
Доцент, кандидат филологических наук,
зав. кафедрой немецкого и французского языков
Петрозаводский государственный университет
E-mail: koturova@petrsu.ru
ORCID ID: 0000-0001-6766-0458
Scopus Author ID: 57699104800

Authors

Elena V. Borzova

(Russia, Petrozavodsk)
Professor, Dr. Sci. (Educ.),
Professor of The English Language Department,
Petrozavodsk State University
E-mail: borzovaelena40@gmail.com
ORCID: 0000-0002-5006-1480
Scopus Author ID: 57210163835

Irina A. Kotiurova

(Russia, Petrozavodsk)
Associate Professor, Cand. Sci. (Philol.),
Head of the Department of German and French Languages
Petrozavodsk State University
E-mail: koturova@petrsu.ru
ORCID ID: 0000-0001-6766-0458
Scopus Author ID: 57699104800

Вклад авторов

Борзова Е. В.: концептуализация, методология, создание рукописи и её редактирование
Котюрова И. А.: концептуализация, методология, проведение исследования, создание рукописи и её редактирование

Authors contribution

Elena V. Borzova: Conceptualization, Methodology, Writing - Review & Editing
Irina A. Kotiurova: Conceptualization, Methodology, Investigation, Writing - Review & Editing

© Борзова Е. В., Котюрова И. А., 2026

Авторы заявляют об отсутствии конфликта интересов

© Elena V. Borzova, Irina A. Kotiurova, 2026

The authors declared no conflicts of interest