



Castledown

 OPEN ACCESS

Technology in Language Teaching & Learning

ISSN 2652-1687

<https://www.castledown.com/journals/tltl/>

Technology in Language Teaching & Learning, 7(1), 102690 (2025)

<https://doi.org/10.29140/tltl.v7n1.102690>

A Critical Review of SkELL (Sketch Engine for Language Learning) – V1.11.8



IBRAHIM HALIL TOPAL 

Gazi University, Türkiye

ibrahimtopal@gazi.edu.tr

Abstract

In an era of technology transforming education, data-driven learning (DDL) has gained recognition as a powerful approach in language education. One way to harness technology in language education is through online resources, one example of which is SkELL (Sketch Engine for Language Learning) (Baisa & Suchomel, 2014). Sketch Engine's simplified language learning interface offers learners authentic usage of words and phrases by tapping into its mother corpus query and management system's large corpora. Designed as a user-friendly alternative to traditional corpus tools, SkELL bridges the gap between complex linguistic analysis and practical language learning. Embracing Topal's (2022) framework, this media review critically evaluates SkELL by addressing its strengths and weaknesses as a language learning resource. Findings reveal that the platform provides authentic usage examples, vocabulary gains, collaborative DDL tasks, and writing improvement. Nonetheless, it suffers from several drawbacks, including the requirement for advanced language use, a lack of differentiation between homographs and polysemous words, and inaccurate/misleading input. Consequently, more research is urged to verify the review's findings.

Keywords: corpus, data-driven learning, language learning, media review, SkELL

Copyright: © 2025 Ibrahim Halil Topal. This is an open access article distributed under the terms of the Creative Commons Attribution Non-Commercial 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. **Data Availability Statement:** All relevant data are within this paper.

Introduction

Traditionally, most language learning is guided by teachers and textbooks. However, more contemporary approaches, especially technological advancements, have facilitated language learning and offered valuable benefits for language learners and teachers (Su & Zou, 2022). One such approach to foreign language learning is data-driven learning (DDL), which treats language learners as active participants in language learning rather than passive recipients (Pérez-Paredes, 2025). Equally, it considers language as data and learners as linguistic researchers (Martins et al., 2025). Learners harnessing DDL engage with authentic language input to discover usage patterns, collocations, and grammatical structures (Crosthwaite, 2024). Grounded in corpus linguistics, DDL draws heavily on bodies of texts, also known as corpora, to provide empirical evidence of authentic language use (Boulton & Cobb, 2017). Rather than prescriptive rules, DDL allows for heuristic and inductive means of language learning and supports autonomous learning (He & Xie, 2024).

The advancement of computer technologies has contributed considerably to the significant place corpora occupy in language education, including authentic language usage, vocabulary, grammar, and writing enhancement (Crosthwaite, 2024; Vyatkina, 2020). However, traditional corpus tools might pose some challenges for language learners, such as complex use, requirement for specialized knowledge, and inaccessibility to learners without computational linguistics expertise (Anthony, 2024). This gap necessitates user-friendly tools that balance computational power with pedagogical accessibility, eventually making corpus-based learning insightful and practical for language learners and teachers. To achieve this, educational innovations must be grounded in both learning theory and data analytics to ensure that tools are not only technologically advanced but also equitable and pedagogically sound (Bernacki, 2025).

Sketch Engine for Language Learning (SkELL) emerges as a solution to these challenges. Unlike conventional corpus tools (e.g., XAIRA, BNCWeb, and AntConc), SkELL is designed specifically for language learners, providing concordancing, word sketches (collocations), and thesaurus (synonyms) in a simple interface (Baisa & Suchomel, 2014). With these features, corpus tools like SkELL bridge the gap between corpus linguistics and classroom learning, democratizes access to linguistic data, and allows language learners to interact with authentic language examples without prior specialized corpus knowledge (Kilgariff et al., 2015; Pérez-Paredes, 2024).

All things considered, this review evaluates SkELL in terms of its potential for language learning and teaching. The justifications for this review are threefold. First, SkELL epitomizes DDL by empowering learners to discover authentic language patterns. Second, SkELL stands out by balancing simplicity with functionality, while other tools (e.g., AntConc and Sketch Engine) require specialized knowledge of corpus management and linguistics. Third, the existing related literature is scant and differs in focus and scope. For instance, Baisa and Suchomel (2014) introduced the tool by characterizing its general features. Similarly, Kilgariff et al. (2015) explored its classroom use along with Sketch Engine. Hirata and Hirata (2018) explored its use among Japanese EFL learners and reported freedom to access large language patterns. Equally, Troy (2022) examined SkELL usage among Japanese EFL learners and noted learner enjoyment and vocabulary improvement. Also, Ahmad (2023) conducted a comparative study of two corpus tools (Playphrase.me and SkELL) for perceiving phrasal verbs and idioms.

Employing Topal's (2022) framework, the review centers around the general description, technical abilities, pedagogical affordances, and a conclusion. Hence, SkELL will initially be described succinctly. The tool's technical and pedagogical qualities will then be characterized. The review will next end with a conclusion and recommendations for future research.

General Description

SkELL is a simplified version of Sketch Engine (<https://www.sketchengine.eu/>), developed by researchers from the Lexical Computing company. It is specifically designed for language learners and teachers to discover authentic language use through the Sketch Engine's large database. SkELL is easily accessible at <https://skell.sketchengine.eu/>. It welcomes users with a ready-made search engine and exempts registration or login. Using SkELL is straightforward, with a main search bar (Figure 1) in the middle and two tabs (in the form of question and exclamation marks). The search bar allows users to perform queries in Czech (csSKELL), English, Estonian (etSKELL), German (deSKELL), Italian (itSKELL), and Russian (ruSKELL).

The results are presented in contextualized examples (concordancing), word sketches (collocations), and synonyms (thesaurus) (Figure 2).

The question mark directs users toward further information about SkELL, while the exclamation mark allows them to report a problem. Tabs for additional settings (ad-free version, terms of use, about SkELL, and cookies) are available at the bottom right page.

Technical Features

User Interface and Navigation

With a clean and straightforward interface, the platform requires no knowledge of corpus linguistics. Users can access three main functions (i.e., examples, word sketch, and similar words) with a single click. Unlike many corpus tools, SkELL is freely available without login barriers. However, SkELL offers limited customization. That is, it does not allow users to adjust settings, such as corpus selection or advanced search preferences. Similarly, it does not offer tutorials or ongoing guides. Hence, first-time users may still benefit from a guided introduction while the platform is intuitive.

Search and Filtering

The simple one-box search system enables users to obtain relevant results instantly by inputting any word or phrase. This simplified search does not require Boolean operators or linguistic tags. The automatic filtering options allow users to receive learner-friendly examples, grammatically organized and



Figure 1 Main Search Bar.

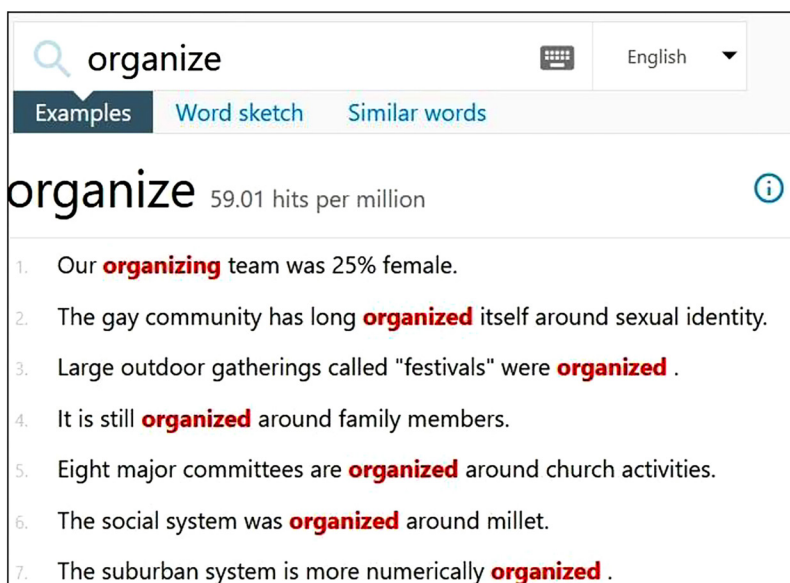


Figure 2 *The Results Presented As Examples (Concordancing), Word Sketch (Collocations), and Similar Words (Thesaurus).*

accurate collocations, synonyms, and semantically related words. Nevertheless, SkELL does not offer advanced filtering options. Therefore, users cannot specify corpus domains, registers, or English varieties. While single-word/phrase search works well, multi-word phrase search may be less effective.

Database and Content Quality

The high-quality, pre-filtered content allows users to retrieve grammatically correct examples. SkELL also utilizes large, reputable corpora (e.g., the TenTen corpora), a family of corpora created from the web, with a corpus size of 10+ billion words per language (currently available in 50+ languages). This massive, web-based corpus ensures extensive and updated language data. In addition, SkELL helps learners move beyond textbook examples to real-world language use by focusing on natural, authentic usage. Furthermore, SkELL offers content in six languages (Czech, English, Estonian, German, Italian, and Russian).

Technical Compatibility

SkELL is browser-based and mobile-friendly, working well on desktops, tablets, and smartphones. It also does not require software installation. Fast processing provides almost instant results. The platform runs smoothly on any device with an internet connection. However, this also means that it has no offline mode. Also, it cannot be linked to other educational platforms or tools for further analysis because of limited API integration.

Pedagogical Affordances

Authentic Input in Multiple Languages

One of SkELL's distinctive features is lexical queries (concordancing, collocations, and thesaurus) in six different languages. Based on their query items, language users can view forty examples in context (Figure 3), correct word combinations (Figure 4), and similar words in word clouds (Figure 5). This feature, in fact, has several pedagogical affordances. In the first place, SkELL provides

portare 580.14 hits per million

1. Il cambiamento scaturisce sempre da qualcosa, e **porta** qualcosa.
2. Nel frattempo **portate** una pentola piena d' acqua ad ebollizione e ricordate di salare l'acqua.
3. Successivamente, è stato **portato** in una luminosa città di cristallo.
4. A volte la cosa migliore che puoi **portare** a casa dal viaggio è la pelle intatta.
5. La registrazione di un nuovo dominio **porta** spesso a numerosi dubbi su quale estensione scegliere.
6. Ogni esperienza **porta** contatti, e magari anche possibilità inaspettate.
7. La donna ha il vantaggio di poter **portare** i capelli lunghi.
8. Durante la rappresentazione, il primo attore ha avuto un malore **portando** comunque a termine lo spettacolo.
9. Proprio l'amore per la terra **porta** a non dimenticare mai il rispetto per la natura.
10. Gli studenti partecipanti sono invitati a **portare** la ricevuta nel giorno del seminario.
11. Allarme **porta** aperta, allarme ottico e acustico per l'aumento della temperatura.
12. Una trattoria di quartiere, tanto da **portare** il nome della chiesa che ha di fronte.
13. Gli scavi **portarono** alla luce un articolato insediamento, raccolto intorno ad un mastio difensivo centrale.
14. La liberazione dei corpi è l'unica strada che può **portarci** alla vera libertà.
15. Certamente la fame è uno degli stimoli più forti che **porta** un neonato a chiamare.
16. Apparentemente, sul mercato, non esiste ancora alcun altro ristorante che **porta** il medesimo nome.
17. Il poggiatesta in particolare delinea molte caratteristiche del possessore che lo **porta** sempre con sé.

Figure 3 Contextualized Examples in Italian (Concordancing).

Frühstück noun ✓ Show context

modifiers of "Frühstück"	nouns with "Frühstück" as genitive object	genitive objects of "Frühstück"
1. reichhaltig einem reichhaltigen Frühstück	1. Buffet	1. gibts Frühstück gibts
2. ausgiebig Nach einem ausgiebigen Frühstück	2. Zubereitung der Zubereitung des Frühstücks	2. Extraklasse
3. lecker einem leckeren Frühstück	3. Zubereiten	3. Tag das Frühstück die wichtigste Mahlzeit des Tages
4. kontinental ein kontinentales Frühstück	4. Einnahme	4. Leben Frühstück Deines Lebens
5. gesund ein gesundes Frühstück	5. Vorbereitung Vorbereitung des Frühstücks	5. Kind
6. gemeinsam einem gemeinsamen Frühstück	6. Bestandteil Bestandteil des Frühstücks	6. Welt
7. ausgewogen ein ausgewogenes Frühstück	7. Genuss	7. Stadt
8. gemütlich einem gemütlichen Frühstück	8. Organisation	
9. ausgedehnt einem ausgedehnten Frühstück	9. Form	
10. herzhaft einem herzhaften Frühstück	10. Rest	
11. üppig einem üppigen Frühstück	11. Rahmen	
12. stärkend Nach einem stärkenden Frühstück	12. Qualität	
13. morgendlich das morgendliche Frühstück	13. Beginn	
14. köstlich einem köstlichen Frühstück	14. Teil	
15. kräftig Nach einem kräftigen Frühstück	15. Tag	

Figure 4 Word Sketch in German (Collocational Search).



Figure 5 Search for Similar Words in English (Thesaurus).

users with authentic usage examples since it draws on an extensive real-life body of texts. Research has demonstrated that multiple corpus examples foster language comprehension and production (Frankenberg-Garcia, 2014).

Interactive Exploration of Collocations and Synonyms

One further benefit of word sketches is that language users can click on hypertexts to display query items collocated by specific words. For instance, adverbs modified by “hence” reveal 15 results, one of which is “ultimately.” When users click on this word, they are directed toward examples of “hence + ultimately” (six results). Similarly, similar words are presented in word clouds. Bigger font size indicates close similarity and vice versa. Additionally, hypertexts of synonyms are also given at the top page under “similar words.”

Contextualized Grammar

In addition to vocabulary improvement, SkELL exposes users to contextualized grammar. For example, searching for prepositions or modal verbs generates phrases, such as *on both sides*, and *one might question*, helping learners understand grammatical structures in context. Research supports SkELL’s role in grammar learning and teaching, highlighting corpora as valuable tools for linguistic development (Rodríguez-Fuentes & Swatek, 2022).

Versatility in Classroom and Self-Study Contexts

Another strength of SkELL is its convenience for classroom and self-study contexts. While teachers can design collaborative DDL tasks, such as student groups working to analyze possessive adjectives and pronouns, independent language users can opt to explore idiomatic expressions or subtle differences between synonyms. This cross-media suitability was also highlighted in earlier scholarly work (Charles, 2023), supporting our argument.

Enhancing Writing Through Data-Driven Learning

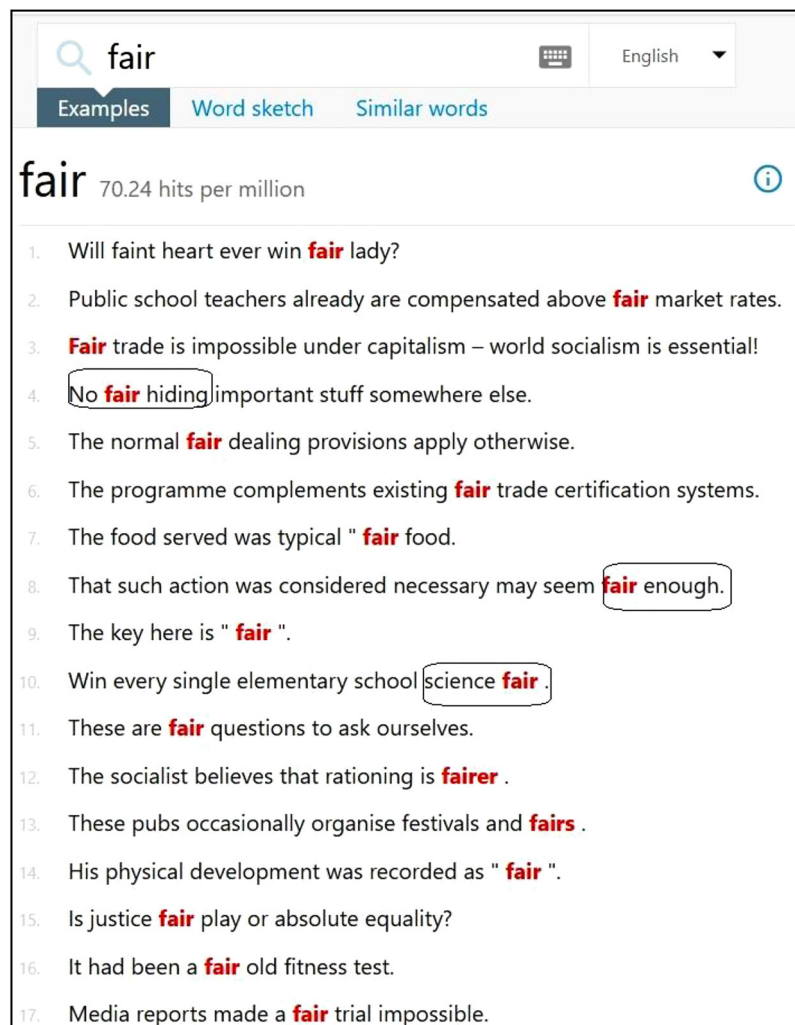
A third benefit of SkELL pertains to writing improvement. More specifically, it enables unsure users while drafting an essay to see the contextual differences in, say, discourse markers (e.g., *in conclusion*

vs. *as a conclusion*), thereby revealing the prior's significantly more prevalence in academic contexts than the latter's. Previous studies also showed DDL's effectiveness in enhancing second language writing performances (Crosthwaite & Steeples, 2024; Luo, 2016), substantiating our claim.

Challenges and Limitations of SkELL

Notwithstanding its affordances, SkELL suffers from several potential pitfalls. The biggest drawback is possibly that it requires advanced language use or addresses autonomous language users. Correspondingly, the lack of interactive feedback might complicate the corpus-driven learning experience for beginning-level students as they might need scaffolding or explanations for collocational usage.

A second disadvantage is that SkELL does not differentiate between homographs or polysemous words. As Figure 4 illustrates, SkELL does not provide additional explanation or have any filters regarding homonymy or polysemy, resulting in scrambled words, meanings, and senses. As Figure 6 indicates, the polysemous word "fair" was not filtered by meaning or part of speech, causing potential confusion among learners.



Search results for "fair" (70.24 hits per million):

- Will faint heart ever win **fair** lady?
- Public school teachers already are compensated above **fair** market rates.
- Fair** trade is impossible under capitalism – world socialism is essential!
- No **fair** hiding important stuff somewhere else.
- The normal **fair** dealing provisions apply otherwise.
- The programme complements existing **fair** trade certification systems.
- The food served was typical " **fair** food.
- That such action was considered necessary may seem **fair** enough.
- The key here is " **fair** ".
- Win every single elementary school science **fair** .
- These are **fair** questions to ask ourselves.
- The socialist believes that rationing is **fairer** .
- These pubs occasionally organise festivals and **fairs** .
- His physical development was recorded as " **fair** ".
- Is justice **fair** play or absolute equality?
- It had been a **fair** old fitness test.
- Media reports made a **fair** trial impossible.

Figure 6 Problems with Homographs and Polysemous Words.

In conclusion English Examples Word sketch Similar words		
conclusion noun ✓ Show context		
verbs with "conclusion" as subject	verbs with conclusion as object	adjective predicates of "conclusion"
1. regard conclusions regarding the	1. draw draw conclusions	1. inescapable The conclusion is inescapable
2. concern conclusions concerning the	2. reach reached the conclusion	2. false conclusion is false .
3. contradict conclusions contradict	3. support support the conclusion	3. premature conclusion is premature ,
4. rest conclusion rests on	4. base conclusions based on	4. speculative conclusions are speculative
5. seem conclusion seems to	5. justify justify the conclusion	5. obvious conclusion is obvious
6. mislead conclusion almost certainly misled	6. near neared its conclusion	6. correct conclusions are correct
7. reach conclusions reached .	7. dispute disputed this conclusion	7. erroneous conclusion is erroneous
8. reinforce a conclusion reinforced by the	8. derive conclusions derived	8. unavoidable conclusion is unavoidable :
9. follow conclusion follows	9. forgo a forgone conclusion .	9. tentative conclusion is only tentative
10. emerge conclusions emerge	10. confirm conclusion is confirmed	10. logical conclusions logical
11. apply conclusion applies	11. yield yields the conclusion	11. contrary conclusion is contrary to
12. surprise surprised by two conclusions	12. summarize summarizing the conclusions	12. true conclusion is true
13. let . In conclusion let	13. contradict contradict this conclusion	13. probable the conclusion probable (
14. differ conclusions differ from	14. reinforce reinforce the conclusion	14. incorrect conclusions are incorrect
15. ignore conclusion ignores	15. compel compel the conclusion that	15. startling conclusions are startling .

Figure 7 An Example of Misleading Output.

Another downside of SkELL usage is providing inaccurate or misleading output (Figure 7). While searching for discourse markers (e.g., *in conclusion*), users might experience difficulty identifying near-synonyms due to wrong or missing outputs. The search may sometimes yield no results for other lexical items (e.g., *consequently*) with close synonyms (e.g., *as a result* and *accordingly*) due to case-insensitive queries (*Consequently* vs. *consequently*).

Conclusion

In conclusion, SkELL is a flourishing corpus web interface applicable to foreign language education. Its main pedagogical affordances are multilingual support in key phrase search, contextualized vocabulary and grammar, writing improvement, and curricular and extracurricular use. However, the review also voiced some concerns regarding its use, including the requirement for advanced/autonomous language use, absence of interactivity, problems with homonymy and polysemy, and inaccurate/misleading results. Nonetheless, SkELL is a promising DDL tool addressing language learners and teachers. Further research, especially in languages other than English, is necessary to corroborate the review's findings about SkELL and reveal its further potential across diverse contexts.

As an EFL teacher in Türkiye, I have found that SkELL is a useful, albeit somewhat sufficient, tool for implementing corpus-based learning in my classes. Because of its basic platform and direct access to authentic data, I can use it in activities designed to build vocabulary, collocation, and grammar skills. I usually create guided discovery tasks where students work with word sketches or example sentences with the goal of finding out how they are used in real life. Yet, some of my lower proficiency learners sometimes have a hard time interpreting the results on their own because SkELL lacks interactive feedback or explanations. This means I must scaffold the process by offering extra assistance, such as guiding students through the analysis or pre-teaching critical concepts. Also, the inability to focus

on academic English or exam English because of the supplied corpora is a challenge for many of my students. Even with all these limitations, I believe SkELL continues to be a valuable additional source for fostering learner autonomy and authentic language use.

References

- Ahmad, A. F. (2023). *Examining the efficiency of PlayPhrase. me and SKELL in supporting the understanding of phrasal verbs and idioms by adult learners of English* (Unpublished Master's thesis). Mid Sweden University, Sweden. <https://miun.diva-portal.org/smash/record.jsf?pid=diva2%3A1806244>
- Anthony, L. (2024). Addressing the challenges of data-driven learning through corpus tool design – In conversation with Laurence Anthony. In P. Crosthwaite (Ed.), *Corpora for language learning: Bridging the research-practice divide* (pp. 9–24). Routledge.
- Baisa, V. & Suchomel, V. (2014). SkELL: Web interface for English language learning. In A. Horák & P. Rychlý (Eds.), *Eighth workshop on recent advances in Slavonic natural language processing* (pp. 63–70). Tribun EU.
- Bernacki, M. L. (2025). Leveraging learning theory and analytics to produce grounded, innovative, data-driven, equitable improvements to teaching and learning. *Journal of Educational Psychology*, 117(1), 1–11. <https://psycnet.apa.org/doi/10.1037/edu0000933>
- Boulton, A., & Cobb, T. (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2), 348–393. <https://doi.org/10.1111/lang.12224>
- Charles, M. (2023). Corpora and autonomous language learning. In R. Jablonkai & E. Csomay (Eds.), *The Routledge handbook of corpora and English language teaching and learning* (pp. 406–419). Routledge.
- Crosthwaite, P. (2024). *Corpora for language learning: Bridging the research-practice divide*. Routledge.
- Crosthwaite, P., & Steeples, B. (2024). Data-driven learning with younger learners: Exploring corpus-assisted development of the passive voice for science writing with female secondary school students. *Computer Assisted Language Learning*, 37(5–6), 1166–1197. <https://doi.org/10.1080/09588221.2022.2068615>
- Frankenberg-Garcia, A. (2014). The use of corpus examples for language comprehension and production. *ReCALL*, 26(2), 128–146. <https://doi.org/10.1017/S0958344014000093>
- He, M., & Xie, Q. (2024). Empowering autonomy in language learning: the sustainable impact of data-driven learning on noun collocation acquisition. *Humanities and Social Sciences Communications*, 11(1), 1–14. <https://doi.org/10.1057/s41599-024-04038-6>
- Hirata, Y., & Hirata, Y. (2018). Students' evaluation of SkELL: The 'Sketch Engine for Language Learning'. In S.K.S. Cheung, L. Kwok, K. Kubota, L-K. Lee, & J. Tokito (Eds.), *Blended learning. Enhancing learning success: 11th International Conference, ICBL 2018, Osaka, Japan, July 31-August 2, 2018, Proceedings 11* (pp. 368–377). Springer.
- Kilgariff, A., Marcowicz, F., Smith, S., & Thomas, J. (2015). Corpora and language learning with the Sketch Engine and SKELL. *Revue Française de Linguistique Appliquée*, 20(1), 61–80. <https://doi.org/10.3917/rfla.201.0061>
- Kohnke, L., & Moorhouse, B. L. (2022). 'There's an app for that!' Writing Technology reviews for academic journals. *RELC Journal*, 53(1), 261–265. <https://doi.org/10.1177/0033688220945419>
- Luo, Q. (2016). The effects of data-driven learning activities on EFL learners' writing development. *SpringerPlus*, 5, 1–13. <https://doi.org/10.1186/s40064-016-2935-5>
- Martins, A. C., Silva, I. R., Marques, J. G. T., & De Martino, J. M. (2025). Teaching sign language as second language through data-driven learning. In D. Li & J. Corbett (Eds.), *The Routledge handbook of corpus translation studies* (pp. 401–416). Routledge.

- Pérez-Paredes, P. (2024). Data-driven learning in informal contexts? Embracing Broad Data-driven learning (BDDL) research. In P. Crosthwaite (Ed.). *Corpora for language learning: Bridging the research-practice divide* (pp. 211–221). Routledge.
- Pérez-Paredes, P. (2025). MALL and data-driven learning. In L. McCallum & D. Tafazoli (Eds.), *The Palgrave encyclopedia of computer-assisted language learning* (pp. 1–7). Springer Nature Switzerland.
- Rodríguez-Fuentes, R. A., & Swatek, A. M. (2022). Exploring the effect of corpus-informed and conventional homework materials on fostering EFL students' grammatical construction learning. *System*, 104, 102676. <https://doi.org/10.1016/j.system.2021.102676>
- Su, F., & Zou, D. (2022). Technology-enhanced collaborative language learning: theoretical foundations, technologies, and implications. *Computer Assisted Language Learning*, 35(8), 1754–1788. <https://doi.org/10.1080/09588221.2020.1831545>
- Troy, H. R. (2022). Learning vocabulary with SkELL: Developing a methodology with university students in Japan using action research. *Bulletin of Nagoya University of Foreign Studies*, 10, 163–183. <http://id.ndl.go.jp/bib/032107918>
- Vyatkina, N. (2020). Corpora as open educational resources for language teaching. *Foreign Language Annals*, 53(2), 359–370. <https://doi.org/10.1111/flan.12464>

Appendix

The Product at a Glance

Title	SKELL (Sketch Engine for Language Learning)
Developer	Lexical Computing Ltd.
Website	https://skell.sketchengine.eu/
Product type	Web-based corpus tool for language learning
Supported languages	Czech, English, Estonian, German, Italian, and Russian
Desktop operating systems	Latest versions of Windows, macOS, and Linux
Mobile operating systems	Latest versions of iOS and Android
Registration	No registration required
Hardware requirements	Any device with an internet connection and a modern web browser
Price	Free of charge