*Proceedings of the 13th Conference on Language Resources and Evaluation (LREC 2022)*, pages 5580–5586
Marseille, 20-25 June 2022


# The Hebrew Essay Corpus

**Chen Gafni, Anat Prior, Shuly Wintner**
University of Haifa, Israel
chen.gafni@gmail.com, aprior@edu.haifa.ac.il, shuly@cs.haifa.ac.il

**Abstract**

We present the Hebrew Essay Corpus: an annotated corpus of Hebrew language argumentative essays authored by prospective higher-education students. The corpus includes both essays by native speakers, written as part of the psychometric exam that is used to assess their future success in academic studies; and essays authored by non-native speakers, with three different native languages, that were written as part of a language aptitude test. The corpus is uniformly encoded and stored. The non-native essays were annotated with target hypotheses whose main goal is to make the texts amenable to automatic processing (morphological and syntactic analysis). The corpus is available for academic purposes upon request. We describe the corpus and the error correction and annotation schemes used in its analysis. In addition to introducing this new resource, we discuss the challenges of identifying and analyzing non-native language use in general, and propose various ways for dealing with these challenges.

**Keywords:** Learner corpora, Hebrew, non-native language

## 1. Introduction

*Learner corpora*—the systematic collection of spoken or written language produced by learners of a language—have been used in research since the late 1980's (Granger, 2002; Tono, 2003; Granger et al., 2015; De Knop and Meunier, 2015). They can follow different designs, be of different sizes, involve different language pairs, etc.[1] One paradigm in analyzing learner corpora is the quantitative comparison of categories (words, multi-word expressions, parts of speech, etc.) between learner corpora and native speaker corpora (Granger, 1996; Gilquin, 2008; Granger, 2015). This approach, which we follow here, is often called *Contrastive Interlanguage Analysis*. The quantitative analyses range from descriptive comparisons (overuse/underuse studies (Durrant and Schmitt, 2009; Gilquin and Paquot, 2008; Hirschmann et al., 2013)) to more involved statistical methods, up to modeling (Gries, 2008; Gries, 2015; Gries and Deshors, 2015; Vyatkina et al., 2015).

Learner and other non-native language corpora have been shown to be instrumental in several tasks, including automatic detection of highly competent non-native writers (Tomokiyo and Jones, 2001; Estival et al., 2007; Bergsma et al., 2012), identification of learners' native language (Koppel et al., 2005; Bykh and Meurers, 2012; Tetreault et al., 2013; Tsvetkov et al., 2013; Goldin et al., 2018) and typology-driven error prediction in learners' language production (Berzak et al., 2015).

In this paper we present the Hebrew Essay Corpus: an annotated corpus of Hebrew language argumentative essays authored by prospective students in higher-education. The corpus includes both essays by native (or near-native) speakers, written as part of a college entry exam that is used to assess their future success in academic studies; and essays authored by non-native speakers, with three different native languages, written as part of a language aptitude test. The corpus is uniformly encoded and stored. The non-native essays were annotated with target hypotheses (Reznicek et al., 2013) whose main goal is to make the texts amenable to automatic processing (morphological and syntactic analysis), thereby guaranteeing uniform representation and processing of the entire dataset.

This paper thus makes two main contributions. The more specific one is the introduction of the Hebrew Essay Corpus. More generally, we propose guidelines and recommendations for meaningful linguistic analysis of non-native texts, which take into account the inherent variability of language, with a focus on Hebrew as the target language. The corpus documentation includes guidelines for specific issues in non-native Hebrew, intended to standardize the analysis as much as possible. In addition, it includes general guidelines intended to increase the awareness of annotators to the issue of linguistic variability.

## 2. The corpus

### 2.1. The essays

The corpus includes 3000 argumentative essays authored by non-native speakers of Hebrew, distributed equally over three native languages (L1s; 1000 essays per L1): Arabic, French, and Russian. In addition, it also includes 1000 essays in Hebrew authored by native speakers. The essays in both collections were written by examinees as part of the admission process to higher education institutions in Israel. The essays by Hebrew native speakers were written as part of the Psychometric Test, a general test required for admission by most higher education programs in Israel. The authors were either native speakers of Hebrew or candidates who decided to take the psychometric test in Hebrew even though they were not native speakers (the

[1] For a list of learner corpora, see Learner Corpora around the World; for an extensive bibliography covering learner corpus analyses, see the resources page of the Learner Corpus Association.

test is also administered in several other languages).[2] The essays by non-native speakers were collected as part of the *YAEL* test: a Hebrew proficiency test required for examinees who chose to sit the Psychometric Test in a language other than Hebrew. Both tests are administered by the Israeli National Institute for Testing and Evaluation (*NITE*), from which we obtained the essays. Essays in the YAEL sub-corpus were written in response to one of nine prompts, while essays in the Psychometric sub-corpus were written about one of two topics (the prompts for the two sub-corpora differ). The psychometric (native) essays were collected in 2012 (topic 1) and 2017 (topic 2). The YAEL (non-native) essays were collected between the years 2011-2020.

## 2.2. Metadata

Essays in the non-native sub-corpus are accompanied by the following metadata (some pieces of information are unavailable for some essays):

**Author's L1** Arabic, French, or Russian.

**Sex** Male, Female, Unspecified.

**Age** 13-50 (average: 21).

**Year of exam** 2011-2020.

**Prompt** 1–9, representing the topic of the essay (the explicit prompts are confidential).

**Essay score** the range of scores for essays included in the corpus is 17-28 (average: 20.7). These scores were assigned by two professional NITE raters.

**Scores of components of essay evaluation** these include (i) Content, (ii) Organization, (iii) Linguistic Richness, and (iv) Linguistic Precision. The range of each component grade is 1-7.

**Total Psychometric score** the scores of the Psychometric test have a normal distribution in the range 200-800 with a mean of 550. The Psychometric scores of candidates whose essays are included in our corpus were in the range 279-778 (average: 540).

**Scores of Psychometric components** (i) Verbal Reasoning, (ii) Quantitative Reasoning, and (iii) English. The range of each component is 50-150.

**Parental education (for each parent):** no education, primary, partial secondary, full secondary, partial tertiary, academic degrees: bachelor, master, doctoral.

**Family income** six levels ranging from very low to very high, plus unspecified income.

The only available metadata for the native speaker essays is the essay score, in the range 1–6 (average: 3.67).

Table 1 summarizes the average number of sentences and tokens per essay for each of the three L1s. The average number of sentences per essay in the native sub-corpus was 15.2 (SD: 5.3), and the average number of tokens was 329 (SD: 81). These numbers are considerably higher than in the non-native essays. However, it is important to note that the allotted time for essay writing was 15 minutes in the YAEL test and 30 minutes in the psychometric test. In addition, there was a specific length requirement for each test: 10-15 lines in YAEL, and 25 lines in the psychometric test. Thus, the length differences across the two sub-corpora are likely due to the test requirements.

| | **Arabic** | **French** | **Russian** |
|---|---|---|---|
| Sentences | 6.1 (2.6) | 9.0 (2.9) | 8.9 (2.7) |
| Tokens | 143 (28) | 142 (29) | 138 (27) |

Table 1: Average number of sentences and tokens per essay across L1s in the YAEL corpus. Numbers in parentheses denote standard deviation.

Figure 1 shows the distribution of essays in the non-native sub-corpus by score. The distribution is evidently normal, but its lower (left) part is truncated by design: we requested only essays above a certain score, because the level of the Hebrew in the lowest-scored essays was too low. Scores can be non-integral because they represent the average of the two human-assigned scores.

Figure 2 depicts the average number of sentences (represented as bars) and tokens (represented as a curve) per essay across the non-native test scores. The number of tokens is significantly correlated with the test score (Pearson's $r = 0.31$, $p < 0.001$), while the number of sentences is not (Pearson's $r = 0.03$, $p = 0.16$).

## 2.3. Processing

The essays, originally hand-written, were transcribed by NITE and stored in text files. The order of sentences in each essay was scrambled before the files were delivered to us to preserve author privacy. Half of the YAEL essays were further reviewed and corrected by NITE for various errors (orthographic, morphological, etc.)

We tokenized the entire dataset using Child Phonology Analyzer (Gafni, 2015), and used our own scripts to generate token-matched original and revised versions for essays corrected by NITE. This included specific marking of tokens that were inserted or deleted in the NITE revised (or corrected) version. We then conducted manual correction and annotation of a subset of the texts, according to principles outlined in Section 3. The modified version of the corpus is stored in a similar way, with indications of additions, deletions and, sometimes, splitting of one token into multiple tokens and vice versa.

We then added morphological and syntactic information about each token to the corpus, including part-of-speech and dependency relations. We combined the output of the YAP parser (More et al., 2019) and the information extracted with NITE's in-house automatic tools. The information obtained from both tools was reviewed and corrected manually, where necessary.

[2] We did not obtain information on the native language of these authors, and we therefore consider them all "natives".

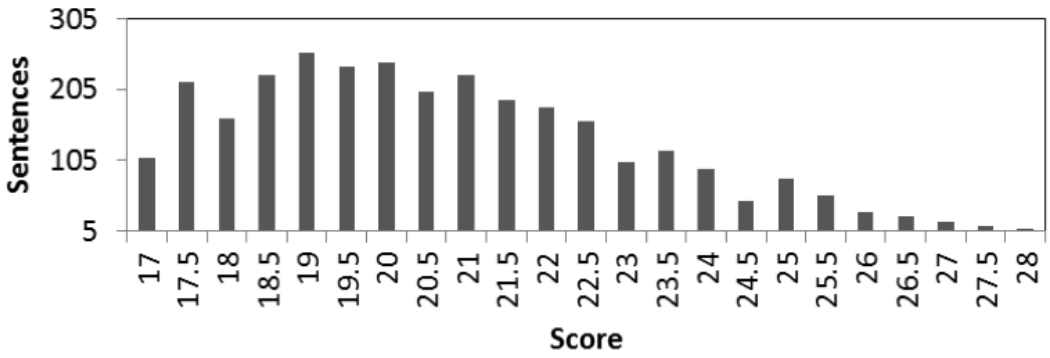


Figure 1: Distribution of essays by score.

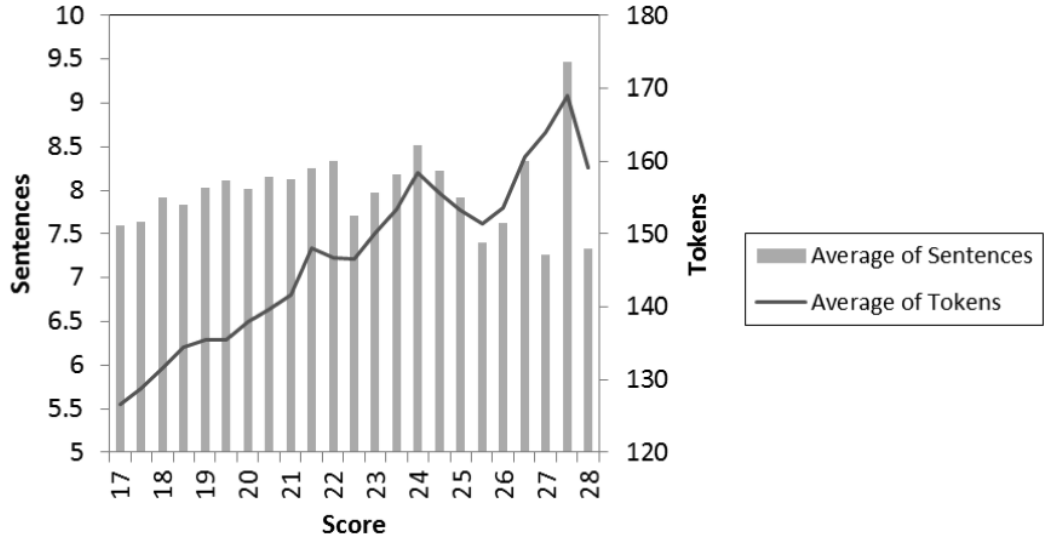


Figure 2: Average number of tokens and sentences per essay across YAEL test scores.

All the datasets will eventually be made publicly available for research purposes in the ANNIS search tool (Krause and Zeldes, 2014; Krause et al., 2017).

## 3. Annotation

We reviewed the essays and corrected various types of errors. In essays that were previously reviewed by NITE, we relied on the existing corrections, but occasionally made additional corrections or revised existing corrections (we also retained NITE's original revision). Most essays were reviewed by one annotator, except for fifty essays that were reviewed by two annotators to assess inter-annotator agreement (see Section 3.3). All annotators were native speakers of Hebrew, with an undergraduate or a graduate degree in linguistics. The remainder of this section details our annotation schema. So far, we completed the annotation of approximately 1000 essays (out of the 3000 non-native ones; we focused first on those that were not corrected by NITE), but we intend to manually annotate the entire corpus.

### 3.1. Principles

When correcting a non-native text, it is sometimes assumed that the language used deviates in some way from "typical", or "standard" native language, and that the author's intended meaning can be recovered and reconstructed according to the norms of the target language. In reality, this is not a straightforward matter. First, the notion of a "standard" native language is elusive: native speakers vary greatly in their use of language, and more often than not avoid adhering to prescriptive language norms (Dabrowska, 2018). Second, non-native texts may contain linguistic expressions that

are either ambiguous or unintelligible. In such cases, it is impossible to construct with certainty an utterance in native-like language that would retain the author's intended meaning, simply because this meaning is not part of the text and is thus unknown.

Therefore, generating an equivalent "native-like" version of a non-native text is a difficult, ill-defined task. Instead, we adopt an approach that minimally modifies the non-native texts by associating some (ill-formed) constructions with a *target hypothesis* (Reznicek et al., 2013). Our goal is to introduce a minimal number of changes in an input sentence in order to obtain a morphologically and syntactically correct utterance in the target language, that would make the resulting utterance amenable to automatic language processing tools, such as a morphological analyzer and a parser.

Having said that, we do extend the scope of our corrections to certain (mostly, lexical) modifications that are not strictly necessary for automatic processing. The motivation behind this decision stems from the nature of the Hebrew language, and in particular its complex morphology and deficient orthography (Fabri et al., 2014). The paucity of vowels in the Hebrew script, combined with the rich morphology of the language and the tendency of some particles (the definite article, the coordinating and subordinating conjunctions, and four highly frequent prepositions) to attach in the script to the word that immediately follows them, result in a situation whereby the "orthographic neighborhood" of a word is unusually large. In other words, the chances of a single spelling error (omission, addition, or substitution of a single letter, or swapping of two adjacent characters) to result in a valid lexical form are much higher than in other languages. Therefore, we took the liberty of correcting some obvious errors that we attribute to such causes.

### 3.2. Examples

We list below some example utterances from the dataset, along with our correction, to illustrate the points discussed above.

First, we correct spelling errors that result in non-words, as in (1):

(1) *harbe talmidim maxlitim lamšix*
*many student.PL decide.M.PL.PRS ?*
*et ha-limudim*
*ACC the-studies*
Presumably: 'Many students decide to continue studying'

The word form למשיך *'lamšix'* does not exist in Hebrew; we correct it to the similar להמשיך *'lehamšix, to continue'*, which results in a grammatical sentence.

To further illustrate the principle of correcting to obtain a grammatical sentence, consider (2):

(2) *ze mašehu roce ba-xaim šelo*
*this something want.SG.PRS in.the-life his*
*'This is something wants in his life'

(2) is ungrammatical. The most conservative interpretation would be to treat משהו *mašehu 'something'* as a morpho-orthographic error, an incorrect merging of the words מה שהוא *ma še-hu 'what that-he'*. The hypothesized target sentence is then (3):

(3) *ze ma šehu roce ba-xaim šelo*
*this what that-he want.SG.PRS in.the-life his*
'This is what he wants in his life'

We therefore annotate the original sentence by indicating the error and its correction.

Sometimes, however, we introduce corrections also to grammatical sentences. Consider (4):

(4) *ze yavo rak le-tola'at*
*it come.3SG.M.FUT only to-worm.CONSTR*
*ha-mišpaxa acma*
*the-family herself*
'It will come only to the worm of the family itself'

Sentence (4) is syntactically correct, but does not make sense in the context in which it appeared. A plausible explanation for this sentence is a letter transposition: לתולעת *'to the worm of'* should probably have been לתועלת *'to the benefit of'*. We annotate this as a spelling error and introduce a correction.

When two alternative corrections are possible, we implement the one requiring minimal assumptions and minimal modifications of the original text. The following examples illustrate this principle.

(5) *lehaspik lahem et kol*
*suffice.INF to.them ACC all*
*corxeyhem*
*needs.POSS.3PL.M*
*'To suffice them all their needs'

(5) is ungrammatical due to a mismatch between the verb and its arguments. The verb להספיק *'suffice'* is assigned two internal arguments here: *[lahem] 'to them'* and *[et kol corxeyhem] 'all their needs'*. Of the two arguments, only the first fits into the argument structure of the verb. However, omitting the second argument will lead to loss of information.

The more plausible correction involves changing the verb להספיק */lehaSPiK/ 'suffice'* to a verb of the same root in a different *binyan* (verb pattern): לספק */leSaPeK/ 'to provide'*. The revised verb is compatible with the argument structure of the original sentence. Thus, no information is lost in the revised sentence and the correction requires a single morphological change. The hypothesized target phrase is then:

(6) *lesapek lahem et kol*
*provide.INF to.them ACC all*
*corxeyhem*
*needs.POSS.3PL.M*
'To provide them all their needs'

In many cases, the author expresses an idea in a way that is atypical of native language, and there is some uncertainty about the intended meaning or the appropriate correction. If the intended meaning seems clear, but there are several equivalent corrections, we choose one of the alternatives and add comments regarding the other alternatives. However, if the intended meaning is not clear, or if there are several plausible corrections that require different syntactic structures, we do not correct the text. Instead, we highlight the problematic text, make verbal comments, or assign a special error tag to parts of the text during the error annotation process.

The detailed annotation schema, along with the guidelines to the annotators, will be made publicly available when they stabilize.

### 3.3. Evaluation

Due to the complexity of the annotation process, the notion of inter-annotator agreement became complex as well. We calculated inter-annotator agreement on several levels: (i) whether annotators agreed that some word or expression contained an error, (ii) whether they applied the same correction, and (iii) whether they annotated the error similarly (even when the correction is identical, the error can be attributed to different sources, such as spelling or morphology). Yet, while identical annotations can be easily labeled as "agreement" among annotators, cases of non-identical annotation do not necessarily mean that annotators truly disagreed. The most obvious counterexample is annotator errors, such as failing to notice an error in the text, and mislabeling errors. Such errors obviously can be tracked and resolved in texts analyzed by multiple annotators, however, there is no guarantee that texts analyzed by a single annotator will not contain annotation errors.

We chose 54 essays, at various proficiency levels and across all three L1s, to be annotated and corrected by two experienced annotators. In total, this evaluation set included 428 sentences comprising 7764 tokens. It is interesting to note that the number of words corrected by both annotators was 786, over 10% of all tokens. Looking only at the binary question, whether both annotators treated word tokens in the same way (i.e., left untouched or corrected), the agreement between the two annotators was 93.5%. A harsher measure looks at the proportion of tokens that were corrected identically by both annotators. This measure takes into account (in other words, penalizes disagreement on) both the binary decision (whether to correct a token) and the actual correction. Here, since the annotators had more freedom in determining the "intended" form of an erroneous token, the agreement was only 50%. All cases of disagreement were then resolved by consultation with a third annotator.

### 3.4. Use case

As a further validation of the quality of the corpus, Nguyen and Wintner (2022) conducted some basic classification experiments. They were able to demonstrate that simple, feature-based classifiers can accurately distinguish between the native and the non-native authors; predict the native language of non-native ones; and quite accurately predict the non-natives' Hebrew proficiency scores. Their models' predictions were often indistinguishable from those of human raters. These results corroborate the hypothesis that there are strong, identifiable signals of the L1 and the authors' proficiency level in the corpus. We therefore trust that the Hebrew Essay Corpus will be invaluable both for research in learner language, transfer effects from L1, etc., and for practical educational applications.

## 4. Conclusion

We presented the Hebrew Essay Corpus, a dataset of essays authored by native and non-native speakers of Hebrew. The dataset was computationally processed, is uniformly represented, and is undergoing error annotation. We expect it to be a valuable resource for any investigation of Hebrew as a second language, specifically when transfer effects from Arabic, French, and Russian are concerned. The corpus, the tools used for processing it, the annotation schema and the guidelines to the annotators are all available for research proposes.

## 5. Acknowledgements

We are immensely grateful to the Israeli National Institute for Testing and Evaluation for making the essays available. We are extremely grateful to Noam Ordan, Anke Lüdeling, Sarah Schneider, Isabelle Nguyen, and Dominique Bobeck for advice and fruitful discussions. This work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 398186468 and by the Data Science Research Center at the University of Haifa.

## 6. Bibliographical References

Bergsma, S., Post, M., and Yarowsky, D. (2012). Stylometric analysis of scientific articles. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 327–337. Association for Computational Linguistics.

Berzak, Y., Reichart, R., and Katz, B. (2015). Contrastive analysis with predictive power: Typology driven estimation of grammatical error distributions in ESL. In *Proceedings of the 19th Conference on Computational Natural Language Learning*, pages 94–102, July.

Bykh, S. and Meurers, D. (2012). Native language identification using recurring $n$-grams – investigating abstraction and domain dependence. In *Proceedings of COLING 2012*, pages 425–440, Mumbai, India, December.

Dabrowska, E. (2018). Experience, aptitude and individual differences in native language ultimate attainment. *Cognition*, 178:222–235.

De Knop, S. and Meunier, F. (2015). The 'learner corpus research, cognitive linguistics and second language acquisition' nexus: a SWOT analysis. *Corpus Linguistics and Linguistic Theory*, 11(1):1–18.

Durrant, P. and Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics*, 47(2):157–177.

Estival, D., Gaustad, T., Pham, S. B., Radford, W., and Hutchinson, B. (2007). Author profiling for English emails. In *Proceedings of the 10th Conference of the Pacific Association for Computational Linguistics*, pages 263–272.

Fabri, R., Gasser, M., Habash, N., Kiraz, G., and Wintner, S. (2014). Linguistic introduction: The orthography, morphology and syntax of Semitic languages. In Imed Zitouni, editor, *Natural Language Processing of Semitic Languages*, Theory and Applications of Natural Language Processing, pages 3–41. Springer, Berlin Heidelberg.

Gafni, C. (2015). Child phonology analyzer: processing and analyzing transcribed speech. In *Proceedings of the 18th International Congress of Phonetic Sciences*, pages 1–5. The Scottish Consortium for ICPhS.

Gilquin, G. and Paquot, M. (2008). Too chatty: Learner academic writing and register variation. *English Text Construction*, 1(1):41–61.

Gilquin, G. (2008). Combining contrastive and interlanguage analysis to apprehend transfer: detection, explanation, evaluation. In Gaëtanelle Gilquin, et al., editors, *Linking Up Contrastive and Learner Corpus Research*, volume 66 of *Language and computers*, pages 3–33. Rodopi, Amsterdam.

Goldin, G., Rabinovich, E., and Wintner, S. (2018). Native language identification with user generated content. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 3591–3601. Association for Computational Linguistics.

Sylviane Granger, et al., editors. (2015). *The Cambridge Handbook of Learner Corpus Research*. Cambridge University Press, Cambridge.

Granger, S. (1996). From CA to CIA and back: An integrated approach to computerized bilingual and learner corpora. In K. Aijmer, et al., editors, *Languages in contrast. Text-based cross-linguistic studies*, pages 37–51. Lund University Press, Lund.

Granger, S. (2002). A bird's-eye view of learner corpus research. In Sylviane Granger, editor, *Computer learner corpora, second language acquisition, and foreign language teaching*, volume 6 of *Language learning and language teaching*, pages 3–33. Benjamins, Amsterdam.

Granger, S. (2015). Contrastive interlanguage analysis: A reappraisal. *International Journal of Learner Corpus Research*, 1(1):7–24.

Gries, S. T. and Deshors, S. C. (2015). EFL and/vs. ESL. a multi-level regression modelling on bridging the paradigm gap. *International Journal of Learner Corpus Research*, 1(1):130–159.

Gries, S. T. (2008). Corpus-based methods in analyses of second language acquisition data. In Peter Robinson et al., editors, *Handbook of Cognitive Linguistics and Second Language Acquisition*, pages 406–431. Routledge.

Gries, S. T. (2015). Statistics for learner corpus research. In Sylviane Granger, et al., editors, *The Cambridge Handbook of Learner Corpus Research*, pages 159–181. Cambridge University Press, Cambridge.

Hirschmann, H., Lüdeling, A., Rehbein, I., Reznicek, M., and Zeldes, A. (2013). Underuse of syntactic categories in Falko. a case study on modification. In Sylviane Granger, et al., editors, *20 Years of Learner Corpus Research. Looking Back, Moving Ahead.*, pages 223–234. Presses Universitaires de Louvain, Louvain la Neuve.

Koppel, M., Schler, J., and Zigdon, K. (2005). Determining an author's native language by mining a text for errors. In *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 624–628. ACM.

Krause, T. and Zeldes, A. (2014). ANNIS3: A new architecture for generic corpus query and visualization. *Digital Scholarship in the Humanities*.

Krause, T., Leser, U., and Lüdeling, A. (2017). graphANNIS: A fast query engine for deeply annotated linguistic corpora. *Journal for Language Technology and Computational Linguistics. Special Issue on Korpuslinguistische Softwarewerkzeuge*, 31(1):1–25.

More, A., Seker, A., Basmova, V., and Tsarfaty, R. (2019). Joint transition-based models for morpho-syntactic parsing: Parsing strategies for MRLs and a case study from Modern Hebrew. *Transactions of the Association for Computational Linguistics*, 7:33–48, March.

Nguyen, I. and Wintner, S. (2022). Predicting the proficiency level of nonnative Hebrew authors. In *Proceedings of LREC-2022*, June.

Reznicek, M., Lüdeling, A., and Hirschmann, H. (2013). Competing target hypotheses in the Falko corpus: A flexible multi-layer corpus architecture. In Ana Díaz-Negrillo, et al., editors, *Automatic Treatment and Analysis of Learner Corpus Data*, pages 101–124. John Benjamins, Amsterdam.

Tetreault, J., Blanchard, D., and Cahill, A. (2013). A report on the first native language identification shared task. In *Proceedings of the Eighth Workshop on Building Educational Applications Using NLP*. Association for Computational Linguistics, June.

Tomokiyo, L. M. and Jones, R. (2001). You're not

from 'round here, are you?: naive Bayes detection of non-native utterance text. In *Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics*, pages 1–8. Association for Computational Linguistics.

Tono, Y. (2003). Learner corpora: Design, development and applications. In Dawn Archer, et al., editors, *Proceedings of the Corpus Linguistics 2003 Conference*, pages 800–809, Lancaster University. University Centre for Computer Corpus Research on Language.

Tsvetkov, Y., Twitto, N., Schneider, N., Ordan, N., Faruqui, M., Chahuneau, V., Wintner, S., and Dyer, C. (2013). Identifying the L1 of non-native writers: the CMU-Haifa system. In *Proceedings of the Eighth Workshop on Innovative Use of NLP for Building Educational Applications*, pages 279–287. Association for Computational Linguistics, June.

Vyatkina, N., Hirschmann, H., and Golcher, F. (2015). Syntactic modification at early stagesof L2 German writing development: A longitudinal learner corpus study. *Journal of Second Language Writing*, 29:28–50, September.