

Automatic word stress annotation of Russian unrestricted text

Robert Reynolds

HSL Faculty

UiT The Arctic University of Norway

N-9018 Tromsø

robert.reynolds@uit.no

Francis Tyers

HSL Faculty

UiT The Arctic University of Norway

N-9018 Tromsø

francis.tyers@uit.no

Abstract

We evaluate the effectiveness of finite-state tools we developed for automatically annotating word stress in Russian unrestricted text. This task is relevant for computer-assisted language learning and text-to-speech. To our knowledge, this is the first study to empirically evaluate the results of this task. Given an adequate lexicon with specified stress, the primary obstacle for correct stress placement is disambiguating homographic wordforms. The baseline performance of this task is 90.07%, (known words only, no morphosyntactic disambiguation). Using a constraint grammar to disambiguate homographs, we achieve 93.21% accuracy with minimal errors. For applications with a higher threshold for errors, we achieved 96.15% accuracy by incorporating frequency-based guessing and a simple algorithm for guessing the stress position on unknown words. These results highlight the need for morphosyntactic disambiguation in the word stress placement task for Russian, and set a standard for future research on this task.

1 Introduction

Lexical stress and its attendant vowel reduction are a prominent feature of spoken Russian; the incorrect placement of stress can render speech almost incomprehensible. This is because Russian word stress is phonemic, i.e. many wordforms are distinguished from one another only by stress position. This is the cause of considerable difficulty for learners, since the inflecting word classes include complex patterns of shifting stress, and a lexeme's stress pattern cannot be predicted from surface forms. Furthermore, standard written Rus-

sian does not typically mark word stress.¹ Without information about lexical stress position, correctly converting written Russian text to speech is impossible. Half of the vowel letters in Russian change their pronunciation significantly, depending on their position relative to the stress. For example the word *dogovórom* 'contract.SG-INS' is pronounced /dɔgɔvɔrɔm/, with the letter *o* realized as three different vowel sounds. Determining these vowel qualities is impossible without specifying the stress position. This is a problem both for humans (e.g. foreign language students) and computers (e.g. text-to-speech).

We identify three different types relations between word stress ambiguity and morphosyntactic ambiguity. First, *intraparadigmatic* stress ambiguity refers to homographic wordforms belonging to the same lexeme, as shown in (1).²

(1) Intraparadigmatic homographs

- a. *téla* 'body.SG-GEN'
- b. *telá* 'body.PL-NOM'

The remaining two types of stress ambiguity occur between lexemes. *Morphosyntactically incongruent* stress ambiguity occurs between homographs that belong to separate lexemes, and whose morphosyntactic values are different, as shown in (2).

(2) Morphosyntactically incongruent homographs

- a. *nášej* 'our.F-SG-GEN/LOC/DAT/INS'
- našej* 'sew on.IMP-2SG'
- b. *doróga* 'road.N-F-SG-NOM'
- dorogá* 'dear.ADJ-F-SG-PRED'

¹Texts intended for native speakers sometimes mark stress on words that cannot be disambiguated through context. Theoretically, a perfect word stress placement system could help an author identify tokens which should be stressed for natives: any token that cannot be disambiguated by syntactic or semantic means should be marked for stress.

²Throughout this article, cyrillic is transliterated using the scientific transliteration scheme.

Morphosyntactically congruent stress ambiguity occurs between homographs that belong to separate lexemes, and whose morphosyntactic values are identical, as shown in (3). This kind of stress ambiguity is relatively rare, and resolving this ambiguity would require the use of technologies such as word sense disambiguation.

- (3) Morphosyntactically congruent homographs
- a. *zámok* ‘castle.SG-NOM’
zamók ‘lock.SG-NOM’
 - b. *zámkov* ‘castle.PL-GEN’
zamkóv ‘lock.PL-GEN’
 - c. ...
 - ...

It should be noted that most morphosyntactic ambiguity in unrestricted text does not result in stress ambiguity. For example, *novyj* ‘new’ (and every other adjective) has identical forms for F-SG-GEN, F-SG-LOC, F-SG-DAT and F-SG-INS: *nóvoj*. Likewise, the form *vypej* has multiple possible readings (including ‘drink.IMP’, ‘bittern.PL-GEN’), but they all have the same stress position: *výpej*. We refer to this as *stress-irrelevant* morphosyntactic ambiguity, since all readings have the same stress placement.

In the case of unrestricted text in Russian, most stress placement ambiguity is rooted in intraparadigmatic and morphosyntactically incongruent ambiguity. Detailed part-of-speech tagging with morphosyntactic analysis can help determine the stress of these forms, since each alternative stress placement is tied to a different tag sequence. In this study we focus on the role of detailed part-of-speech tagging in improving automatic stress placement. We leave morphosyntactically congruent stress ambiguity to future work because it is by far the least common type of stress ambiguity (less than 1% of tokens in running text), and disambiguating morphosyntactically congruent stress requires fundamentally different technology from the other approaches of this study.

1.1 Background and task definition

Automatic stress placement in Russian is similar to diacritic restoration, a task which has received increasing interest over the last 20 years. Generally speaking, diacritics disambiguate otherwise homographic wordforms, so missing diacritics can complicate many NLP tasks, such as

text-to-speech. For example, speakers of Czech may type emails and other communications without standard diacritics. In order to generate speech from these texts, they must first be normalized by restoring diacritics. A slightly different situation arises with languages whose standard orthography is underspecified, like vowel quality in Arabic or Hebrew. For such languages, the ‘restoration’ of vowel diacritics results in less ambiguity than in standard orthography. For languages with inherently ambiguous orthography, it may be more precise to refer to this as ‘diacritic enhancement’, since it produces text that is less ambiguous than the standard language. In this sense, Russian orthography is similar to Arabic and Hebrew, since its vowel qualities are underspecified in standard orthography.

Many studies of Russian text-to-speech and automatic speech recognition make note of the difficulties caused by the shortcomings of their stress-marking resources (e.g. Krivnova (1998)). Text-to-speech technology must deal with the inherent ambiguity of Russian stress placement, and many articles mention disambiguation of one kind or another, but to our knowledge no studies have empirically evaluated the success of their approach. Several studies have investigated methods for predicting stress position on unknown words. For example, Xomicevič et al. (2008) developed a set of heuristics for guessing stress placement on unknown words in Russian. More recently, Hall and Sproat (2013) trained a maximum entropy model on a dictionary of Russian words, and evaluated on wordlists containing ‘known’ and ‘unknown’ wordforms.³ Their model achieved 98.7% accuracy on known words, and 83.9% accuracy on unknown words. The task of training and evaluating on wordlists is different from that of placing stress in running text. Since many of the most problematic stress ambiguities in Russian occur in high-frequency wordforms, evaluations of wordform lists encounter stress ambiguity seven times less frequently than in running text (see discussion in Section 4). Furthermore, working with running text includes the possibility of disambiguating homographs based on syntactic context.

³Hall and Sproat (2013) randomly selected their training and test data from a list of wordforms, and so a number of lexemes had wordforms in both the training and test data. Wordforms in the test data whose sibling wordforms from the same lexeme were in the training set were categorized as ‘known’ wordforms.

So far, the implicit target application of the few studies related to automatic stress placement in Russian has been text-to-speech and automatic speech recognition. However, the target application of our stress annotator is in a different domain: language learning. Since standard Russian does not mark word stress, learners are frequently unable to pronounce unknown words correctly without referencing a dictionary or similar resources. In the context of language learning, marking stress incorrectly is arguably worse than not marking it at all. Because of this, we want our stress annotator to be able to abstain from marking stress on words that it is unable to resolve with high confidence.

1.2 Stress corpus

Russian texts with marked word stress are relatively rare, except in materials for second language learners, which are predominantly proprietary. Our gold-standard corpus was collected from free texts on Russian language-learning websites. This small corpus (7689 tokens) is representative of texts that learners of Russian are likely to encounter in their studies. These texts include excerpts from well-known literary works, as well as dialogs, prose, and individual sentences that were written for learners.

Unfortunately, the general practice for marking stress in Russian is to *not* mark stress on monosyllabic tokens, effectively assuming that all monosyllabics are stressed. However, this approach is not well-motivated. Many words – both monosyllabic and multisyllabic – are unstressed, especially among prepositions, conjunctions, and particles. Furthermore, there are many high-frequency monosyllabic homographs that can be either stressed or unstressed, depending on their part of speech, or particular collocations. For example, the token *čto* is stressed when it means ‘what’ and unstressed in the conjunction *potomu čto* ‘because’. For such words, one cannot simply assume that they are stressed on the basis of their syllable count.

Based on these considerations, we built our tools to mark stress on every word, both monosyllabic and multisyllabic. However, because our gold-standard corpus texts do not mark stress on monosyllabic words, we cannot evaluate our annotation of those words.

Similarly, some compound Russian words have

secondary stress, but this is rarely marked, if at all, even in educational materials. Therefore, even though our tools are built to mark secondary stress, we cannot evaluate secondary stress marks, since they are absent in our gold-standard corpus.

In order to test our word stress placement system, we removed all stress marks from the gold-standard corpus, then marked stress on the unstressed version using our tools, and then compared with the original.

2 Automatic stress placement

State-of-the-art morphological analysis in Russian is based on finite-state technology (Nožov, 2003; Segalovich, 2003). To our knowledge, no existing open-source, broad-coverage resources are available for analyzing and generating stressed wordforms. Therefore, we developed free and open-source finite-state tools capable of analyzing and generating stressed wordforms, based on the well-known *Grammatical Dictionary of Russian* (Zaliznjak, 1977). Our Finite-State Transducer⁴ (FST) generates all possible morphosyntactic readings of each wordform, and our Constraint Grammar⁵ (Karlsson, 1990; Karlsson et al., 1995) then removes some readings based on syntactic context. The ultimate success of our stress placement system depends on the performance of the constraint grammar. Ideally, the constraint grammar would successfully remove all but the correct reading for each token, but in practice some tokens still have more than one reading remaining. Therefore, we also evaluate various approaches to deal with the remaining ambiguity, as described below. Table 1 shows two possible sets of readings for the token *kosti*, as well as the output of each approach described below. The first column exhibits stress ambiguity between the noun readings and the imperative verb reading. The second column shows a similar set of readings, after the constraint grammar has removed the imperative verb reading. This results in only stress-irrelevant ambiguity.

The `bare` approach is to not mark stress on words with more than one reading. Since both sets of readings in Table 1 have more than one reading, `bare` does not output a stressed form.

⁴Using two-level morphology (Koskeniemi, 1983; Koskeniemi, 1984), implemented in both `xfst` (Beesley and Karttunen, 2003) and `hfst` (Linden et al., 2011)

⁵Implemented using `vislcg3` constraint grammar parser (<http://beta.visl.sdu.dk/cg3.html>).

Readings:	КОСТЬ-N-F-SG-GEN	<i>kósti</i>	КОСТЬ-N-F-SG-GEN	<i>kósti</i>
	КОСТЬ-N-F-SG-DAT	<i>kósti</i>	КОСТЬ-N-F-SG-DAT	<i>kósti</i>
	КОСТИТЬ-V-IPFV-IMP	<i>kostí</i>		
bare		<i>kosti</i>		<i>kosti</i>
safe		<i>kosti</i>		<i>kósti</i>
randReading		<i>kósti</i> ($p=0.67$) or <i>kostí</i> ($p=0.33$)		<i>kósti</i>
freqReading		<i>kósti</i>		<i>kósti</i>

Table 1: Example output of each stress placement approach, given a particular set of readings for the token *kosti*

The `safe` approach is to mark stress only on tokens whose morphosyntactic ambiguity is stress-irrelevant. In Table 1, the first column has readings that result in two different stress positions, so `safe` does not output a stressed form. However, in the second column, both readings have the same stress position, so `safe` outputs that stress position.

The `randReading` approach is to randomly select one of the available readings. In the first column of Table 1, a random selection means that *kósti* is twice as likely as *kostí*. The second column of Table 1 contains stress-irrelevant ambiguity, so a random selection of a reading has the same result as the `safe` approach.

The `freqReading` approach is to select the reading that is most frequent, with frequency data taken from a separate hand-disambiguated corpus. If none of the readings are found in the corpus, then `freqReading` selects the reading with the tag sequence (lemma removed) that is most frequent in our corpus. If the tag sequence is not found in our frequency list, then `freqReading` backs off to the `randReading` algorithm. In the first column of Table 1, `freqReading` selects *kósti* because the tag sequence N-F-SG-GEN is more frequent than the other alternatives. Note that for tokens with stress-irrelevant ambiguity (e.g. the second column of Table 1), `randReading` and `freqReading` produce the same result as the `safe` method.

So far, the approaches discussed are dependent on the availability of readings from the FST. The focus of our study is on disambiguation of known words, but we also wanted to guess the stress of unknown tokens in order to establish some kind of accuracy maximum for applications that are more tolerant of higher error rates. To this end, we selected a simple guessing method for unknown words. A recent study by Lavitskaya and Kabak (2014) concludes that Russian has default final stress in consonant-final words, and penultimate

stress in vowel-final words.⁶ Based on this conclusion, the `guessSyll` method places the stress on the last vowel that is followed by a consonant.⁷ This method is applied to two unknown wordforms in two approaches, `randReading+guessSyll` and `freqReading+guessSyll`, which are otherwise identical to `randReading` and `freqReading`, respectively.

For our baseline, we take the output of our morphological analyzer (without the constraint grammar) in combination with the `bare`, `safe`, `randReading`, `freqReading`, `randReading+guessSyll`, and `freqReading+guessSyll` approaches. We also compare our outcomes with the RussianGram⁸ plugin for the Google Chrome web browser. RussianGram is not open-source, so we can only guess what technologies support the service. In any case, it provides a meaningful reference point for the success of each of the methods described above.

3 Results

We evaluated all multisyllabic words with marked stress in the gold-standard corpus (N = 4048). Since our approach is lexicon-based, some of our results should be interpreted with respect to how many of the stressed wordforms in the gold-standard corpus can be found in the output of the finite-state transducer. We refer to this measure as *recall*.⁹ Out of 4048 tokens, 3949 were

⁶There is some disagreement over how to define default stress in Russian, cf. Crosswhite et al. (2003).

⁷Although this approach is simplistic, unknown words are not the central focus of this study. More sophisticated heuristics and machine-learning approaches to unknown words are discussed in Section 4.

⁸<http://russiagram.com/>

⁹Our method of computing recall assumes that if even one reading is output by the FST, then all possible readings are present. We have not attempted to formally estimate how frequently this assumption fails, but we expect such cases to be rare.

approach	accuracy%	error%	abstention%	totTry%	totFail%
noCG+bare	30.43	0.17	69.39	30.61	69.57
noCG+safe	90.07	0.49	9.44	90.56	9.93
noCG+randReading	94.34	3.36	2.30	97.70	5.66
noCG+freqReading	95.53	2.59	1.88	98.12	4.47
noCG+randReading+guessSyll	94.99	4.05	0.96	99.04	5.01
noCG+freqReading+guessSyll	95.83	3.46	0.72	99.28	4.17
CG+bare	45.78	0.44	53.78	46.22	54.22
CG+safe	93.21	0.74	6.05	93.95	6.79
CG+randReading	95.50	2.59	1.90	98.10	4.50
CG+freqReading	95.73	2.40	1.88	98.12	4.27
CG+randReading+guessSyll	95.92	3.33	0.74	99.26	4.08
CG+freqReading+guessSyll	96.15	3.14	0.72	99.28	3.85
RussianGram	90.09	0.79	9.12	90.88	9.91

Table 2: Results of stress placement task evaluation

found in the FST, which is equal to 97.55%. This number represents the ceiling for methods relying on the FST. Higher scores are only achievable by expanding the FST’s lexicon or by using syllable-guessing algorithms. After running the constraint grammar, recall was 97.35%, a reduction of 0.20%.

Results were compiled for each of the 13 approaches discussed above: *without* the constraint grammar (noCG) x 6 approaches, *with* the constraint grammar (CG) x 6 approaches, and RussianGram (RussianGram). Results are given in Table 2. Each token was categorized as either an accurate output, or one of two categories of failures: errors and abstentions. If the stress tool outputs a stressed wordform, and it is incorrect, then it is counted as an ‘error’. If the stress tool outputs an unstressed wordform, then it is counted as an ‘abstention’. Abstentions can be result of either unknown wordforms, or known wordforms with no stress specified in our lexicon.

The two right-most columns in Table 2 combine values of the basic categories. The term ‘totTry’ refers to the sum of the accuracy and error rate. This number represents the proportion of tokens on which our system output a stressed wordform. In the case of noCG+bare, the accuracy% (30.43) and error% (0.17) sum to the totTry% value of 30.61. The term ‘totFail’ refers to the sum of error rate and abstention rate, which is the proportion of tokens for which the system failed to output the correct stressed form. In the case of noCG+bare, the error% (0.17) and abstention% (69.39) sum to the totFail% value of 69.57 (rounded).

The noCG+bare approach achieves a baseline accuracy of 30.43%, so roughly two thirds of the tokens in our corpus are morphosyntactically ambiguous. The error rate of 0.17% primarily represents forms whose stress position varies from speaker to speaker (e.g. *zavilis’* vs. *zavilís’* ‘they crinkled’), or errors in the gold-standard corpus (e.g. *verím* ‘we believe’).

The noCG+safe approach achieves a 60% improvement in accuracy (90.07%), which means that 89.39% of morphosyntactic ambiguity on our corpus is stress-irrelevant. Interestingly, the RussianGram web service achieves results that are very close to the noCG+safe approach.

Since the ceiling recall for the FST is 97.55%, and since the noCG+safe approach achieves 90.07%, the maximum improvement that a constraint grammar could theoretically achieve is 7.48%. A comparison of noCG+safe and CG+safe reveals an improvement of 3.14%, which is about 42% of the way to the ceiling recall.

The CG+randReading and CG+freqReading approaches are also limited by the 97.55% ceiling from the FST, and their accuracies achieve improvements of 2.29% and 2.52%, respectively, over CG+safe. However, these gains come at the cost of error rates as much as 3.5 times higher than CG+safe: +1.85% and +1.66%, respectively. It is not surprising that CG+freqReading has higher accuracy and a lower error rate than CG+randReading, since frequency-based guesses are by definition more likely to occur. The frequency data were taken from a very small corpus, and it is likely that frequency from a larger corpus

would yield better results.

The `guessSyll` approach was designed to make a guess on every wordform that is not found in the FST, which would ideally result in an abstention rate of 0%. However, the abstention rates of approximately 0.7% are a manifestation of the fact that some words in the FST, especially proper nouns, have not been assigned stress. Because the FST outputs a form – albeit unstressed – the `guessSyll` algorithm is not called. This means that `guessSyll` is only guessing on about 2% of the tokens. The improvement on overall accuracy from `CG+freqReading` to `CG+freqReading+guessSyll` is 0.42%, which means that the `guessSyll` method guess was accurate 21% of the time.

4 Discussion

One of the main points of this paper is to highlight the importance of syntactic context in the Russian word stress placement task. If your intended application has a low tolerance for error, the `noCG+safe` approach represents the highest accuracy that is possible without leveraging syntactic information for disambiguation (90.07%). In other words, a system that is blind to morphosyntax and contextual disambiguation cannot significantly outperform `noCG+safe`. It would appear that this is the method used by `RussianGram`, since its results are so similar to `noCG+safe`. Indeed, this result can be achieved most efficiently without any part-of-speech tagging, but through simple dictionary lookup.

We noted in Section 1.1 that Hall and Sproat (2013) achieved 98.7% accuracy on stress placement for individual wordforms in a list (i.e. *without* syntax). This result is 8.63% higher than `noCG+safe`, but it is also a fundamentally different task. Based on the surface forms in our FST – which is based on the same dictionary used for Hall and Sproat (2013) – we calculate that only 29 518 (1.05%) of the 2 804 492 wordforms contained in our FST are stress-ambiguous. In our corpus of unrestricted text, at least 7.5% of the tokens are stress-ambiguous. Therefore, stress ambiguity is more than seven times more prevalent in our corpus of unrestricted text than it is in our wordform dictionary. Since the task of word stress placement is virtually always performed on running text, it seems prudent to make use of surrounding contextual information. The experiment

described in this paper demonstrates that a constraint grammar can effectively improve the accuracy of a stress placement system without significantly raising the error rate. Our Russian constraint grammar is under continual development, so we expect higher accuracy in the future.

We are unaware of any other empirical evaluations of Russian word stress placement in unrestricted text. The results of our experiment are promising, but many questions remain unanswered. The experiment was limited by properties of the gold-standard corpus, including its size, genre distribution, and quality. Our gold-standard corpus represents a broad variety of text genres, which makes our results more generalizable, but a larger corpus would allow for evaluating each genre individually. For example, the vast majority of Russian words with shifting stress are of Slavic origins, so we expect a genre such as technical writing to have a lower proportion of words with stress ambiguity, since it contains a higher proportion of borrowed words, calques, and neologisms with simple stress patterns.

In addition to genre, it is also likely that text complexity affects the difficulty of the stress placement task. The distribution of different kinds of syntactic constructions vary with text complexity (Vajjala and Meurers, 2012), and so we expect that the effectiveness of the constraint grammar will be affected by those differences.

The resources needed for machine-learning – such as a large corpus of Russian unrestricted text with marked stress – are simply not available at this time. Even so, lexicon- and rule-based approaches have some advantages over machine-learning approaches. For example, we are able to abstain from marking stress on tokens whose morphosyntactic ambiguity cannot be adequately resolved. In language-learning applications, this reduces the likelihood of learners being exposed to incorrect wordforms, and accepting them as authoritative. Such circumstances can lead to considerable frustration and lack of trust in the learning tool. However, in error-tolerant applications, machine-learning does seem well-suited to placing stress on unknown words, since morphosyntactic analysis is problematic.

The syllable-guessing algorithm `guessSyll` used in this experiment was overly simplistic, and so it was not surprising that it was only moderately successful. More rigorous rule-based approaches

have been suggested in other studies (Church, 1985; Williams, 1987; Xomicevič et al., 2008). For example, Xomicevič et al. (2008) attempt to parse the unknown token by matching known prefixes and suffixes.

Other studies have applied machine-learning to guessing stress of unknown words (Pearson et al., 2000; Webster, 2004; Dou et al., 2009; Hall and Sproat, 2013). For example, Hall and Sproat (2013) achieve an accuracy of 83.9% with unknown words. Their model was trained on a full list of Russian words, which is not representative of the words that would be unknown to a system like ours, so it would be possible modify their approach to fit our application. Most of the complicated word stress patterns are closed classes¹⁰, so we could exclude closed classes of words from the training data, leaving only word classes that are likely to be similar to unknown tokens, such as those with productive derivational affixes.

5 Conclusions

We have demonstrated the effectiveness of using a constraint grammar to improve the results of a Russian word stress placement task in unrestricted text by resolving 42% of the stress ambiguity in our gold-standard corpus. We showed that stress ambiguity is seven times more prevalent in our corpus of running text than it is in our lexicon, suggesting the importance of context-based disambiguation for this task. As with any lexicon- and rule-based system, the lexicon and rules can be expanded and improved, but our initial results are promising, especially considering the short timespan over which they were developed.

As this is the first empirical study of its kind, we also discussed methodological limitations and possibilities for subsequent research, including collecting stressed corpora of varying text complexity and/or genre, as well as implementing and/or adapting established word stress-guessing methods for unknown words.

Acknowledgments

We are very grateful to Laura Janda, Tore Neset and other members of the CLEAR research group at UiT The Arctic University of Norway for comments on an earlier draft of this paper. We

¹⁰The growing number of masculine nouns with shifting stress (*dóktor*~*doktorá* ‘doctor’~‘doctors’) is one exception to this generalization.

are also grateful to three anonymous reviewers for thoughtful criticism. All remaining errors are our own.

References

- Kenneth R. Beesley and Lauri Karttunen. 2003. *Finite State Morphology: Xerox tools and techniques*. CSLI Publications, Stanford.
- Kenneth Church. 1985. Stress assignment in letter to sound rules for speech synthesis. *Association for Computational Linguistics*, pages 246–253.
- Katherine Crosswhite, John Alderete, Tim Beasley, and Vita Markman. 2003. Morphological effects on default stress in novel Russian words. In *WCCFL 22 Proceedings*, pages 151–164.
- Qing Dou, Shane Bergsma, Sittichai Jiampojarn, and Grzegorz Kondrak. 2009. A ranking approach to stress prediction for letter-to-phoneme conversion. In *Proceedings of the Joint Conference of the 47th annual meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*, pages 118–126, Suntec, Singapore. Association for Computational Linguistics.
- Kieth Hall and Richard Sproat. 2013. Russian stress prediction using maximum entropy ranking. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*, pages 879–883, Seattle, Washington, USA. Association for Computational Linguistics.
- Fred Karlsson, Atro Voutilainen, Juha Heikkilä, and Arto Anttila, editors. 1995. *Constraint Grammar: A Language-Independent System for Parsing Unrestricted Text*. Number 4 in Natural Language Processing. Mouton de Gruyter, Berlin and New York.
- Fred Karlsson. 1990. Constraint grammar as a framework for parsing running text. In *Proceedings of the 13th Conference on Computational Linguistics (COLING), Volume 3*, pages 168–173, Helsinki, Finland. Association for Computational Linguistics.
- Kimmo Koskenniemi. 1983. Two-level morphology: A general computational model for word-form recognition and production. Technical report, University of Helsinki, Department of General Linguistics.
- Kimmo Koskenniemi. 1984. A general computational model for word-form recognition and production. In *Proceedings of the 10th International Conference on Computational Linguistics, COLING ’84*, pages 178–181, Stroudsburg, PA, USA. Association for Computational Linguistics.
- Olga F. Krivnova. 1998. Avtomatičeskij sintez ruskoj reči po proizvol’nomu tekstu (vtoraja versija s ženskim golosom) [Automatic Russian speech synthesis with unrestricted text (version 2 with female voice)].

- In *Trudy meždunarodnogo seminara Dialog [Proceedings of the international seminar Dialog]*, pages 498–511.
- Yulia Lavitskaya and Barış Kabak. 2014. Phonological default in the lexical stress system of Russian: Evidence from noun declension. *Lingua*, 150:363–385, Oct.
- Krister Linden, Miikka Silfverberg, Erik Axelson, Sam Hardwick, and Tommi Pirinen. 2011. Hfst—framework for compiling and applying morphologies. In Cerstin Mahlow and Michael Pietrowski, editors, *Systems and Frameworks for Computational Morphology*, volume Vol. 100 of *Communications in Computer and Information Science*, pages 67–85. Springer.
- Igor Nožov. 2003. *Morfologičeskaja i sintaksičeskaja obrabotka teksta (modeli i programmy) [Morphological and Syntactic Text Processing (models and programs)]* also published as *Realizacija avtomatičeskoj sintaksičeskoj segmentacii ruskogo predloženiya [Realization of automatic syntactic segmentation of the Russian sentence]*. Ph.D. thesis, Russian State University for the Humanities, Moscow.
- Steve Pearson, Roland Kuhn, Steven Fincke, and Nick Kibre. 2000. Automatic methods for lexical stress assignment and syllabification. In *International Conference on Spoken Language Processing*, pages 423–426.
- Ilya Segalovich. 2003. A fast morphological algorithm with unknown word guessing induced by a dictionary for a web search engine. In *International Conference on Machine Learning; Models, Technologies and Applications*, pages 273–280.
- Sowmya Vajjala and Detmar Meurers. 2012. On improving the accuracy of readability classification using insights from second language acquisition. In Joel Tetreault, Jill Burstein, and Claudial Leacock, editors, *In Proceedings of the 7th Workshop on Innovative Use of NLP for Building Educational Applications*, pages 163—173, Montréal, Canada, June. Association for Computational Linguistics.
- Gabriel Webster. 2004. Improving letter-to-pronunciation accuracy with automatic morphologically-based stress prediction. In *Eighth International Conference on Spoken Language Processing*, pages 2573–2576.
- Briony Williams. 1987. Word stress assignment in a text-to-speech synthesis system for british english. *Computer Speech and Language*, 2:235–272.
- Olga Xomicević, Sergej Rybin, Andrej Talanov, and Ilya Oparin. 2008. Avtomatičeskoe opredelenie mesta udarenie v neznakomyx slovax v sisteme sinteza reči [Automatic determination of the place of stress in unknown words in a speech synthesis system]. In *Materialy XXXVI meždunarodnoj filologičeskoj konferencii [Proceedings of the XXXVI International Philological Conference]*, Saint Petersburg.
- Andrej Anatoljevič Zaliznjak. 1977. *Grammatičeskij slovar' ruskogo jazyka: slovoizmenenie: okolo 100 000 slov [Grammatical dictionary of the Russian language: Inflection: approx 100 000 words]*. Russkij jazyk.