

Text Modification for Bulgarian Sign Language Users

Slavina Lozanova* Ivelina Stoyanova† Svetlozara Leseva†

slavinal@abv.bg iva@dcl.bas.bg zarka@dcl.bas.bg

Svetla Koeva† Boian Savtchev‡

svetla@dcl.bas.bg bsavtchev@gmail.com

* AssistNet, Sofia, Bulgaria

† Department of Computational Linguistics, IBL, BAS, Sofia, Bulgaria

‡ Cambridge Language Assessment Centre BG015, Sofia, Bulgaria

Abstract

The paper discusses the main issues regarding the reading skills and comprehension proficiency in written Bulgarian of people with communication difficulties, and deaf people, in particular. We consider several key components of text comprehension which pose a challenge for deaf readers and propose a rule-based system for automatic modification of Bulgarian texts intended to facilitate comprehension by deaf people, to assist education, etc. In order to demonstrate the benefits of such a system and to evaluate its performance, we have carried out a study among a group of deaf people who use Bulgarian Sign Language (BulSL) as their primary language (primary BulSL users), which compares the comprehensibility of original texts and their modified versions. The results shows a considerable improvement in readability when using modified texts, but at the same time demonstrates that the level of comprehension is still low, and that a complex set of modifications will have to be implemented to attain satisfactory results.

1 Introduction

The individual development of deaf people depends on a complex of factors, which include the cause and the degree of hearing loss, the age of hearing loss onset, educational background, language and communication methods, cultural identification, disability preconceptions. Hearing loss leads to a limited spoken language input, delays in language acquisition and communication difficulties. Deaf children and adults demonstrate lower reading achievements than hearing people regardless of the degree of hearing loss, and the use (or lack) of high-performing hearing amplifi-

cation devices (Paul, 1998; Conrad, 1979; Musselman, 2000; Traxler, 2000; Vermeulen AM, 2007), which shows that reading skills are influenced by complex social, linguistic and communication-related factors rather than by the sensory disability alone.

The paper explores reading comprehension of Deaf people¹ who use Bulgarian Sign Language (BulSL) as their primary language. Various research studies both in Bulgaria and abroad have shown that hearing-impaired BulSL users have poorer reading skills than their hearing peers. Various methods for text modification have been explored to the end of obtaining texts that correspond to the proficiency of the readers. Most of the modification methodologies have been focused on simplifying the original texts and decreasing their complexity (Inui et al., 2003). Our approach, however, focuses not on simplification, but on the adaptation of the structure of the original texts to the linguistic properties of BulSL.

The paper is organized as follows. Section 2 discusses the reading skills of BulSL users, paying attention to children's and adult education in Bulgaria focused on the acquisition of Bulgarian and the relationship between BulSL and verbal Bulgarian. After outlining the main principles which underlie text adaptation aimed at fostering text comprehensibility in the target population, we present a rule-based method for automatic modification of Bulgarian written texts. The method applies a set of linguistic transformations and produces modified versions of the texts, which are better suited to the needs of BulSL users (Section 3). Section 4 describes an experiment devised to explore the reading comprehension of BulSL users of original and modified texts. Section 5 draws conclusions

¹Capitalized 'Deaf' is used to denote the community of deaf people who use Sign Language as their primary language. The term emphasizes the socio-cultural model of Deafness rather than the medical view of hearing impairment.

and outlines some directions for future work.

2 Reading Skills of Hearing-Impaired People

2.1 Education

Previous research has shown that deaf students lag behind their hearing peers in reading comprehension, because they experience difficulties with vocabulary (Paul, 1996), syntax (Kelly, 1996), and the use of prior knowledge and metacognition (Trezek et al., 2010). In addition, reading comprehension difficulties are often linked to lack of general knowledge due to inadequate education and limited access to information (Lewis and Jackson, 2001). Two independently performed studies (Albertini and Mayer, 2011; Parault and William, 2010) have found out that deaf college students' reading skills are below those of six-graders².

MacAnally et al. (1999) support the hypothesis that using less complicated and more accessible reading materials, consisting of language constructions close or similar to sign language structure can facilitate reading comprehension and motivate deaf people to read. In support of this claim Berent (2004) points out that deaf students would read more smoothly if the subject, verb, and object are in a simple SVO (subject-verb-object) word order. These studies provide evidence in favour of text adaptation that reflects features of the sign language and the development of modified teaching materials.

Bulgarian education for the deaf is based entirely on the oral approach and no systematic effort has been invested into exploring total communication and bilingual approaches (Lozanova, 2002; Saeva, 2010). Even in the specialized schools for the deaf, where sign language communication occurs naturally, BulSL has not been integrated into the school curriculum. Besides, the linguistic analysis of BulSL has been limited to mere descriptions and presentation of signs: Bulgarian Sign Language dictionaries (1966, 1996); Sign Language Dictionary in Civil Education (Stoyanova et al., 2003); Specialized Multimedia BulSL dictionary³ (2005).

In order to improve education and the reading and communication skills of deaf people, a comprehensive study of BulSL is necessary, that will

provide the basis for developing advanced methods for automatic text modification directed to improving text readability for deaf BulSL users.

2.2 Sign Language and Verbal Language

Research has shown that Deaf children of Deaf parents (DCDP) with sign language as their primary mode of communication outperform their deaf peers of hearing parents (DCHP) on different academic tests, including reading tests (Mayberry, 2000). Several studies have found a positive correlation between the advanced American Sign Language (ASL) skills of deaf students and their higher reading skills (Hoffmeister, 2000; Padden and Ramsey, 2000). Evidence is not conclusive as to how sign languages relate to verbal languages and what influence they have on the acquisition of general communication skills and knowledge about the world.

The extensive research on sign languages in the last fifty years worldwide has shown that they are independent linguistic systems which differ from verbal languages (Stokoe, 1960; Stokoe, 1972; Sutton-Spence and Woll, 2003). Being an independent language, a sign language affects the way in which its users conceptualize the world, according to the principle of linguistic relativity, first formulated by Sapir and Whorf (Lee, 1996). Due to the fact that sign languages are very different from verbal ones, many Deaf people attain a certain level of proficiency in a verbal language at the state of interlanguage⁴ (Selinker, 1972) but that level is not sufficient to ensure successful social integration.

2.3 Readability of Written Texts for Native Users of Sign Language

Readability is measured mainly on the basis of vocabulary and sentence complexity, including word length and sentence length: the higher the letter, syllable and word count of linguistic units, the greater the demand on the reader. Some syntactic structures also affect readability – negative and interrogative constructions, passive voice, complex sentences with various relations between the main clause and the subordinates, long distance dependencies, etc. Besides, readability improves if the information in a text is well-organized and effec-

²11-12-year-olds.

³<http://www.signlanguage-bg.com>

⁴The term 'interlanguage' denotes the intermediate state in second language acquisition characterized by insufficient understanding and grammatical and lexical errors in language production.

tively presented so that its local and global discourse structure is obvious to the reader (Swann, 1992).

Text modification is often understood as simplification of text structure but this may result in an inadequately low level of complexity and loss of relevant information. Moreover, using a limited vocabulary, avoiding certain syntactic structures, such as complex sentences, is detrimental to the communication and learning skills.

The efforts towards providing equal access to information for Deaf people lack clear principles and uniformity. Firstly, there is no system of criteria for evaluation of text complexity in terms of vocabulary, syntactic structure, stylistics and pragmatics. Further, no standard framework and requirements for text modification have been established, which limits its applications.

3 Text Modification of Bulgarian

Language modification for improved readability is not a new task and its positive and negative aspects have been extensively discussed (BATOD, 2006). One of the most important arguments against text modification is that it requires a lot of resources in terms of human effort and time. An appealing alternative is to employ NLP methods that will facilitate the implementation of automatic modification for improved readability of written texts aimed at the BulSL community.

3.1 General Principles of Text Modification

Several studies have observed different aspects of text modification: splitting chosen sentences with existing tools (Petersen and Ostendorf, 2007), 'translating' from complex to simplified sentences with statistical machine translation methods (Specia, 2010), developing text simplification systems (Candido et al., 2009), etc. (Siddharthan, 2011) compares a rule-based and a general purpose generator approaches to text adaptation. Recently the availability of the Simple English Wikipedia has provided the opportunity to use purely data-driven approaches (Zhu et al., 2010). The main operation types both in statistical and in rule-based approaches are: *change*, *delete*, *insert*, and *split* (Bott and Saggion, 2011).

Although text modification is a highly language dependent task, it observes certain general principles:

- Modified text should be identical or very

close in meaning to the original.

- Modified text should be grammatically correct and structurally authentic by preserving as much as possible of the original textual and syntactic structure.
- In general, modified text should be characterized by less syntactic complexity compared with the original text. However, the purpose of the modification is not to simplify the text but rather to make the information in it more accessible and understandable by representing it in relatively short information chunks with simple syntax without ellipses.
- It should be possible to extend the range of modifications and include other components which contribute to readability or introduce other functionalities that facilitate reading comprehension, such as visual representations.

3.2 Stages of Text Modification

At present we apply a limited number of modifications: clause splitting, simplification of syntactic structure of complex sentences, anaphora resolution, subject recovery, clause reordering and insertion of additional phrases.

3.2.1 Preprocessing

The preprocessing stage includes annotation with the minimum of grammatical information necessary for the application of the modification rules. The texts are sentence-split, tokenized, POS-tagged and lemmatized using the Bulgarian Language Processing Chain⁵ (Koeva and Genov, 2011). Subsequently, clause splitting is applied using a general method based on POS tagging, lists of clause delimiters – clause linking words and multiword expressions and punctuation, and a set of language specific rules.

We define a clause as a sequence of words between two clause delimiters where exactly one finite verb occurs. A finite verb is either: (a) a single finite verb, e.g. *yade* (*eats*); (b) or a finite verb phrase formed by an auxiliary and a full verb, e.g. *shteshe da yade* (*would eat*); or (c) a finite copular verb phrase with a non-verbal subject complement, e.g. *byaha veseli* (*were merry*).

We identify finite verbs by means of a set of rules applied within a window, currently set up to

⁵<http://dcl.bas.bg/services/>

two words to the left or to the right:

Rule P1. A single finite verb is recognized by the POS tagger. (Some smoothing rules are applied to detect the verb forms actually used in the context – e.g. forms with reflexive and negative particles).

Rule P2. If auxiliaries and a lexical verb form occur within the established window, they form a single finite verb phrase. (This rule subsumes a number of more specific rules that govern the formation of analytical forms of lexical verbs by attaching auxiliary verbs and particles.)

Rule P3. If an auxiliary (or a copular verb) but not a lexical verb form occurs within the established window, the auxiliary or copula itself is a single finite verb.

Rule P4. If a modal and/or a phase verb and a lexical verb form occur within the established window, they form a single finite verb phrase.

Rule P5. If a modal (and/or a phase) verb but not a lexical verb form occurs within the established window, the modal verb itself is a single finite verb.

A clause is labeled by a clause opening (CO) at the beginning and a clause closing (CC) at the end. We assume that at least one clause boundary – an opening and/or a close – occurs between any pair of successive finite verbs in a sentence. Each CO is paired with a CC, even if it might not be expressed by an overt element.

We distinguish two types of COs with respect to the type of clause they introduce: coordinate and subordinate. Most of the coordinating conjunctions in Bulgarian are ambiguous since they can link not only clauses, but also words and phrases. On the contrary, most of the subordinating conjunctions, to the exception of several subordinators which are homonymous with prepositions, particles or adverbs, are unambiguous.

Clause closing delimiters are sentence end, closing *comma*, *colon*, *semicolon*, *dash*.

The following set of clause splitting rules are applied (C1-C9):

Rule C1. The beginning of a sentence is a coordinate CO.

Rule C2. A subordinating clause linking word or phrase denotes a subordinate CO.

Rule C3. If a subordinate CO is on the top of the stack, we look to the right for a punctuation clause delimiter (e.g. *comma*) which functions as a CC element.

Rule C4. If a subordinate CO is on the top of the

stack, and the CC is not identified yet, we look for a coordinating clause linking word or phrase which marks a coordinate CO.

Rule C5. If a coordinate CO is on the top of the stack, we look for another coordinating clause linking word or phrase which marks a coordinate CO.

Rule C6. If a coordinate CO is on the top of the stack and no coordinate CO is found, we look for a punctuation clause delimiter (e.g. a *comma*) which functions as a CC element.

Rule C7. If no clause boundary has been identified between two finite verbs, we insert a clause boundary before the second finite verb.

Rule C8. All COs from the stack should have a corresponding CC.

Rule C9. The part of the sentence to the right of the last finite verb until the end of the sentence should contain the CCs for all COs still in the stack.

3.2.2 Empty subject recovery

The detection, resolution, and assignment of function tags to empty sentence constituents have become subject of interest in relation to parsing (Johnson, 2002; Ryan Gabbard and Marcus, 2004; Dienes and Dubey, 2003), in machine translation, information extraction, automatic summarization (Mitkov, 1999), etc. The inventory of empty categories includes null pronouns, traces of extracted syntactic constituents, empty relative pronouns, etc. So far, we have limited our work to subject recovery.

A common feature of many, if not all, sign languages (BulSL among others) is that each sentence requires an overt subject. Moreover, each subject is indexed by the signer by pointing to the denoted person or thing if it is present in the signing area, or by setting up a point in space as a reference to that person or thing, if it is outside the signing area, and referring to that point whenever the respective person or object is mentioned. In order to avoid ambiguity, different referents are assigned different spatial points. Deaf people find it difficult to deal with complex references in written texts where additional disambiguating markers are rarely available. Being a pro(noun)-drop language, Bulgarian allows the omission of the subject when it is grammatically inferable from the context.

So far the following rules for subject recovery have been defined and implemented:

Rule SR1. In case the verb is in the first or second person singular or plural and the clause lacks a nominative personal pronoun that agrees with the finite verb, a personal pronoun with the respective agreement features is inserted in the text.

Rule SR2. In case the verb is in the third person singular or plural and the clause lacks a noun or a noun phrase that a) precedes the verb; and b) agrees with the verb in person, number and gender, the closest noun (a head in a noun phrase) in the preceding clause that satisfies the agreement features of the verb is inserted in the text. (The precision of the rule for singular verbs is low.)

3.2.3 Anaphora Resolution

With respect to text modification regarding anaphora resolution, we focus on a limited types of pronominal anaphors – personal, relative and possessive pronouns.

Bulgarian personal pronouns agree in gender and number with their antecedent. Possessive pronouns express a relation between a possessor and a possessed item, and agree both with their antecedent (through the root morpheme) and with the head noun (through the number and gender features of the inflection). For instance in the sentence *Vidyah direktora v negovata kola* (*I saw the director in his car*), the possessive pronoun *negov* indicates that the possessor is masculine or neuter singular and the inflection *-a* – that the possessed is feminine gender, singular. The agreement with the possessor is a relevant feature to text modification. Some relative pronouns *koyto* (*which*) (type one) agree with their antecedent in gender and number while others (type two) – *chiyto* (*whose*) agree with the noun they modify and not with their antecedent.

We have formulated the following rules for anaphora resolution:

Rule AR1. The antecedent of a personal or a possessive pronoun is the closest noun (the head in the noun phrase) within a given window to the left of the pronoun which satisfies the agreement features of the pronoun.

Rule AR2. The antecedent of a relative pronoun is the nearest noun (the head in the noun phrase) in the preceding clause that satisfies the agreement features of the pronoun.

The following rules for modification of anaphora can be used:

Rule R1. The third personal pronoun is replaced with the identified antecedent.

Rule R2. The possessive pronoun is replaced with a prepositional phrase formed by the preposition *na* (*of*) and the identified antecedent.

Rule R3. A relative pronoun of type one is replaced with the identified antecedent.

Rule R4. The relative pronoun *chiyto* (*whose*) is replaced with a prepositional phrase formed by the preposition *na* (*of*) and the identified antecedent.

Rule R5. The relative pronoun *kakavto* (*such that*) is replaced by a noun phrase formed by a demonstrative pronoun and the identified antecedent *takava chanta* (*that bag*).

3.2.4 Simplification of Complex Sentences

Complex sentences are one of the main issues for deaf readers because in BulSL, as well as in other sign languages, they are expressed as separate signed statements and the relation between them is explicit.

(Van Valin and LaPolla, 1997) observe that the elements in complex sentences (and other constructions) are linked with a different degree of semantic and syntactic tightness, which is reflected in the Interclausal Relations Hierarchy. The clauses in a sentence have different degree of independence, which determines whether they can be moved within the sentence or whether they can form an individual sentence.

Temporally related events in BulSL most often are represented in a chronological order, and the relation between them is expressed by separate signs or constructions (Example 1).

Example 1.

Zabavlyavayte se, dokato nauchavate i novi neshta.

Have fun while you learn new things.

Signed sentence:

Vie se zabavlyavate. Ednovremenno nauchavate novi neshta /ednovremenno/.

You have fun. Simultaneously, you learn new things /simultaneously/.

(the sign 'simultaneously' can be repeated at the end of the sentence again)

Chambers et al. (2007) and Tatu and Srikanth (2008) identify event attributes and event-event features which are used to describe temporal relations between events. Attributes include tense, grammatical aspect, modality, polarity, event

class. Further, the event-event features include the following: *before*, *includes*, *begins*, *ends*, *simultaneously*, and their respective inverses (Chambers et al., 2007), as well as *sameActor* (binary feature indicating that the events share the same semantic role Agent), *eventCoref* (binary attribute capturing co-reference information), *oneSent* (true when both events are within the same sentence), *relToDocDate* (defining the temporal relation of each event to the document date) (Tatu and Srikanth, 2008).

(Pustejovsky et al., 2003) also introduce temporal functions to capture expressions such as *three years ago*, and use temporal prepositions (*for*, *during*) and temporal connectives (*before*, *while*). Three types of links are considered: TLINK (temporal link between an event and a moment or period of time); SLINK (subordination link between two events); and ALINK (aspectual link between aspectual types).

The structure of the complex sentences is simplified by clause reordering that explicitly reflects the chronological order of the described events. The preposition or postposition of clauses with temporal links *if*, *before*, *after*, etc. may not match the actual causal order. In such cases the order of clauses is simply reversed based on rules of the type:

Temporal link	<i>sled kato /when, after/</i>
Construction	CL1 temporal link CL2
Modification(s)	CL2. <i>Sled tova /then/</i> CL1.

3.2.5 Post-editing

Post editing aims at providing grammatically correct and semantically complete modified text. Clause reordering might lead to inappropriate use of verb tenses. Coping a subject from the previous sentence might require a transformation from an indefinite to a definite noun phrase. Thus, several checks for grammaticality and text cohesion are performed and relevant changes to verb forms and noun definiteness are made. Specific expressions are introduced to highlight temporal, causative, conditional and other relations and to serve as connectives.

Example 2 shows a fully modified text.

Example 2.

Original:

Vaz osnova na doklada ot razsledvaneto, sled kato litseto e bilo uvedomeno za vsichki dokazatelstva i sled kato e bilo izslushano, organat e izdal

razreshenie.

Based on the report from the investigation, after the person has been notified about all evidence and after /he/ has been heard, the authorities have issued a permit.

Modified:

Litseto e bilo uvedomeno za vsichki dokazatelstva. Litseto e bilo izslushano.

Sled tova vaz osnova na doklada ot razsledvaneto, organat mozhe da dade razreshenie.

The person has been notified about all evidence.

The person has been heard.

After that based on the report from the investigation, the authorities may issue a permit.

3.3 Evaluation of System Performance

The evaluation of performance is based on the Bulgarian part of the Bulgarian-English Clause-Aligned Corpus (Koeva et al., 2012) which amounts to 176,397 tokens and includes several categories: administrative texts, fiction, news. The overall evaluation of the system performance is assessed in terms of the evaluation of all subtasks (Section 3.2) as presented in Table 1. The evaluation of finite verbs and anaphora recognition, as well as subject identification is performed manually on a random excerpt of the corpus. Clause splitting is evaluated on the basis of the manual annotation of the corpus. We assess the precision and recall in terms of full recognition and partial recognition. In the first case the entire verb phrase, clause, anaphora, or dropped subject is recognized correctly, while in the latter – only a part of the respective linguistic item is identified. We account for partial recognition since it is often sufficient to produce correct overall results, e.g. partial verb phrase recognition in most cases yields correct clause splitting.

4 Experiments and Evaluation of Readability of Modified Texts

4.1 Outline of the Experiment

4.1.1 Aims and Objectives

The objective of the experiment was to conduct a pilot testing of original and modified texts in order

Task	Precision	Recall	F_1
Finite verb phrases (full)	0.914	0.909	0.912
Finite verb phrases (partial)	0.980	0.975	0.977
Clauses borders	0.806	0.827	0.817
Clauses (beginning)	0.908	0.931	0.919
Anaphora (full)	0.558	0.558	0.558
Anaphora (partial)	0.615	0.615	0.615
Subject (full)	0.590	0.441	0.504
Subject (partial)	0.772	0.548	0.671

Table 1: Evaluation of different stages of text modification

to determine and confirm the need of text modification for deaf people whose primary language is BulSL and the verbal language is acquired as a second language.

The rationale was to identify and distinguish between levels of comprehension of original and automatically modified texts.

4.1.2 Respondents' Profile

The participants were selected regardless of their degree and onset of hearing loss. The experiment targeted the following group of people:

- Socially active adults (18+);
- BulSL users;
- People with developed reading skills.

4.2 Pilot Test Design Methodology and Implementation

4.2.1 Text Selection

We decided to use original and modified versions of journalistic (e.g. news items) and administrative (e.g. legal) texts. The guiding principle was to select texts that are similar in terms of length, complexity, and difficulty.

The selected news refer to topics of general interest such as politics in neighbouring countries, culture, etc. The administrative texts represent real-life scenarios, rather than abstract or rare legal issues. In general, selected texts do not include domain-specific terms and professional jargon.

Regarding text modification the main objective was to preserve the meaning of the original text in

compliance with the principles of textual and factual accuracy and integrity, and appropriate complexity. The result from the automatic modifications has been manually checked and post-edited to ensure grammaticality.

4.2.2 Methodology

The testing is conducted either online via tests in e-form (predominantly), or using paper-based versions. Respondents are given texts of each type, i.e. two original and two modified texts. Each text is associated with two tasks, which have to be completed correctly after the reading. The tasks seek to check the level of understanding of the main idea, details, purpose, implication, temporal relations (the sequence of events), and the ability to follow the text development.

- **Task-type 1: Sequence questions.** The respondents have to arrange text elements (sentences and clauses) listed in a random sequence into a chronological order. The task covers temporal, causative, conditional, and other relations, and its goal is to test reading comprehension which involves temporal and logical relations and inferences.
- **Task-type 2: Multiple response questions (MRQ) for testing general reading comprehension.** MRQ are similar to Multiple choice questions (MCQs) in that they provide a predefined set of options, but MRQ allow any number and combinations of options.

Text	Type	Version	# sentences	# clauses	# temporal shifts
1	News	Original	2	6	2
2	News	Modified	5	6	0
3	Admin	Original	1	4	2
4	Admin	Modified	4	4	0

Table 2: Structure of the test

4.2.3 Structure of the Test

The test consists of four different texts, each of them with two subtasks – for checking the comprehension of temporal relations and the logical structure of the events in the text (type 1), and general comprehension (type 2).

The number of sentences, clauses and temporal shifts for each text is presented in Table 2.

4.3 Analysis of Results

19 deaf adults proficient in BulSL have taken part in the pilot test study. The results are presented in Table 3 and on Figure 1.

Task	Type	Version	correct	all	%
1.1	News	Original	5	19	26.32
2.1	News	Modified	9	19	47.37
3.1	Admin	Original	6	19	31.58
4.1	Admin	Modified	10	19	52.63
1.2	News	Original	7	19	36.84
2.2	News	Modified	9	19	47.37
3.2	Admin	Original	7	19	36.84
4.2	Admin	Modified	10	19	52.63

Table 3: Results of chronological order sub-tasks (1.1-4.1) and general comprehension sub-tasks (1.2-4.2)

We recognize the fact that the small number of respondents does not provide sufficient data to draw conclusions regarding the improvement of readability when using modified texts. However, the results show a significant improvement ($t = 2.0066$ with $p = 0.0485 < 0.05$) in the overall comprehension (chronological order and general understanding) when using the modified texts in comparison with the original texts.

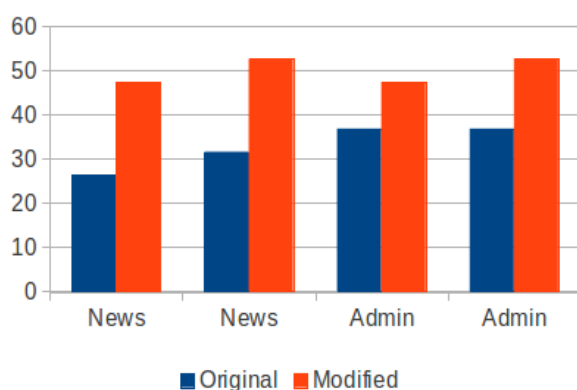


Figure 1: Results in % of correct answers for original and modified texts

Still, the improvement in readability after the text modification is very low and not sufficient to provide reliable communication strategies and access to information. Further work will be aimed at more precise methodology for testing the reading skills of deaf people.

5 Conclusions

As the pilot test suggests, the limited number of modifications is not sufficient to compensate for the problems which deaf people experience with reading. A wider range of text modifications are necessary in order to cover the problematic areas of verbal language competence. Other issues include the use of personal and possessive pronouns, in particular clitics, which are often dropped, the correct use of auxiliary verbs and analytical verb forms. Additional problems such as adjective and noun agreement, subject and verb agreement, etc. need to be addressed specifically, since these have a very different realization in sign languages (e.g., subject and verb are related spatially).

It should be emphasized that there has not been any systematic effort for studying BulSL so far. The detailed exploration of the linguistic properties of BulSL in relation to Bulgarian can give a deeper understanding about the problems in the acquisition of Bulgarian and in particular, the reading difficulties experienced by deaf readers.

Directions for future work include:

- To explore the relationship between reading comprehension and social, educational and other factors;
- To explore the dependence between reading skills and proficiency in BulSL;
- To analyze problems in relation to vocabulary with relation to reading;
- To build a detailed methodology for testing of reading comprehension;
- To explore further the potential of text modification with respect to BulSL in relation to the comparative analyses of the features of BulSL and verbal Bulgarian language.

References

- J. Albertini and C. Mayer. 2011. Using miscue analysis to assess comprehension in deaf college readers. *Journal of Deaf Studies and Deaf Education*, 16:35–46.
- BATOD. 2006. Training materials for language modification.
- Gerald Berent. 2004. Deaf students' command of English relative clauses (Paper presented at the annual convention of Teachers of English to Speakers of Other Languages).

- Stefan Bott and Horacio Saggon. 2011. Spanish text simplification: An exploratory study. In *XXVII Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural (SEPLN 2011)*, Huevla, Spain.
- Arnaldo Candido, Erick Maziero, Caroline Gasperin, Thiago A. S. Pardo, Lucia Specia, and Sandra Aluisio. 2009. Supporting the adaptation of texts for poor literacy readers: a text simplification editor for Brazilian Portuguese. In *Proceedings of the Fourth Workshop on Innovative Use of NLP for Building Educational Applications*, pages 34–42.
- Nathanael Chambers, Shan Wang, and Dan Jurafsky. 2007. Classifying temporal relations between events. In *Proceedings of the ACL 2007 Demo and Poster Sessions, Prague, June 2007*, pages 173–176.
- R. Conrad. 1979. *The deaf schoolchild*. London: Harper and Row.
- Peter Dienes and Amit Dubey. 2003. Deep syntactic processing by combining shallow methods. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, volume 1.
- R. J. Hoffmeister. 2000. A piece of the puzzle: ASL and reading comprehension in deaf children. In C. Chamberlain, J. P. Morford, and R. I. Mayberry, editors, *Language acquisition by eye*, pages 143–163. Mahwah, NJ: Erlbaum.
- Kentaro Inui, Atsushi Fujita, Tetsuro Takahashi, Ryu Iida, and Tomoya Iwakura. 2003. Text simplification for reading assistance: A project note. In *PARAPHRASE '03 Proceedings of the second international workshop on Paraphrasing*, volume 16, pages 9–16. Association for Computational Linguistics Stroudsburg, PA, USA.
- Mark Johnson. 2002. A simple pattern-matching algorithm for recovering empty nodes and their antecedents. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*.
- L. Kelly. 1996. The interaction of syntactic competence and vocabulary during reading by deaf students. *Journal of Deaf Studies and Deaf Education*, 1(1):75–90.
- S. Koeva and A. Genov. 2011. Bulgarian language processing chain. In *Proceedings of Integration of multilingual resources and tools in Web applications. Workshop in conjunction with GSCL 2011, University of Hamburg*.
- Svetla Koeva, Borislav Rizov, Ekaterina Tarpomanova, Tsvetana Dimitrova, Rositsa Dekova, Ivelina Stoyanova, Svetlozara Leseva, Hristina Kukova, and Angel Genov. 2012. Bulgarian-english sentence- and clause-aligned corpus. In *Proceedings of the Second Workshop on Annotation of Corpora for Research in the Humanities (ACRH-2)*, Lisbon, 29 November 2012, pages 51–62. Lisboa: Colibri.
- Penny Lee. 1996. *The Logic and Development of the Linguistic Relativity Principle. The Whorf Theory Complex: A Critical Reconstruction*. John Benjamins Publishing.
- Margaret Jelinek Lewis and Dorothy Jackson. 2001. Television literacy: Comprehension of program content using closed-captions for the deaf. *Journal of Deaf Studies and Deaf Education*, 6(1):43–53.
- Slavina Lozanova. 2002. Predpostavki za razvítie na bilingvizma kato metod za obuchenie na detsa s uvreden sluh. *Spetsialna pedagogika*.
- R. I. Mayberry. 2000. Cognitive development of deaf children: The interface of language and perception in cognitive neuroscience. In *Child Neuropsychology, Volume 7 of handbook of neuropsychology*, pages 71–107. Amsterdam: Elsevier.
- P. McAnally, S. Rose, and S. Quigley. 1999. *Reading practices with deaf learners*. Austin, TX: Pro-Ed.
- Ruslan Mitkov. 1999. Anaphora resolution: the state of the art; working paper, (based on the coling'98/acl'98 tutorial on anaphora resolution).
- Carol Musselman. 2000. How do children who can't hear learn to read an alphabetic script? a review of the literature on reading and deafness. *Journal of Deaf Studies and Deaf Education*, 5:9–31.
- C. Padden and C Ramsey. 2000. American Sign Language and reading ability in deaf children. In C. Chamberlain, J. P. Morford, and R. I. Mayberry, editors, *Language acquisition by eye*, pages 165–189. Mahwah, NJ: Erlbaum.
- S. J. Parault and H. William. 2010. Reading motivation, reading comprehension and reading amount in deaf and hearing adults. *Journal of Deaf Studies and Deaf Education*, 15:120–135.
- Peter Paul. 1996. Reading vocabulary knowledge and deafness. *Journal of Deaf Studies and Deaf Education*, 1(1):3–15.
- Peter Paul. 1998. *Literacy and deafness: The development of reading, writing and literate thought*. Needham, MA: Allyn & Bacon.
- Sarah Petersen and Mari Ostendorf. 2007. *Natural language processing tools for reading level assessment and text simplification for bilingual education*. University of Washington Seattle, WA.
- James Pustejovsky, José Casta no, Robert Ingria, Roser Saurí, Robert Gaizauskas, Andrea Setzer, Graham Katz, and Dragomir Radev. 2003. TimeML: Robust Specification of Event and Temporal Expressions in Text. Technical report, AAAI.
- Seth Kulick Ryan Gabbard and Mitchell Marcus. 2004. Using linguistic principles to recover empty categories. In *Proceedings of the 42nd Annual Meeting on Association For Computational Linguistics*, page 184191.

- S. Saeva. 2010. *Gluhota i Bilingvizam*. Aeropres BG.
- L. Selinker. 1972. Interlanguage. *International Review of Applied Linguistics*, 10:209–231.
- Advait Siddharthan. 2011. Text simplification using typed dependencies: A comparison of the robustness of different generation strategies. In *Proceedings of the 13th European Workshop on Natural Language Generation (ENLG)*, pages 211, September.
- L. Specia. 2010. Translating from complex to simplified sentences. In *9th International Conference on Computational Processing of the Portuguese Language (Propor-2010), Porto Alegre, Brazil*, volume 6001 of *Lecture Notes in Artificial Intelligence*, pages 30–39. Springer.
- William Stokoe. 1960. Sign language structure: An outline of the visual communication systems of the american deaf. *Studies in linguistics: Occasional papers*, 8.
- William Stokoe. 1972. *Semiotics and Human Sign Languages*. NICI, Printers, Ghent.
- Ivelina Stoyanova, Tanya Dimitrova, and Viktorija Trajkovska. 2003. A handbook in civil education with a sign language dictionary. In *Social and Educational Training for Hearing Impaired youths: A Handbook in Civil Education with a Sign Language Dictionary*. Petar Beron, Sofia. (in Bulgarian).
- Rachel Sutton-Spence and Bencie Woll. 2003. *The Linguistics of British Sign Language: An Introduction*. Cambridge University Press, 3rd edition.
- W. Swann. 1992. *Learning for All: Classroom Diversity*. Milton Keynes: The Open University.
- Marta Tatu and Munirathnam Srikanth. 2008. Experiments with reasoning for temporal relations between events. In *COLING '08 Proceedings of the 22nd International Conference on Computational Linguistics*, volume 1, pages 857–864.
- C. B. Traxler. 2000. The stanford achievement test, 9th edition: National norming and performance standards for deaf and hard-of-hearing students. *Journal of Deaf Studies and Deaf Education*, 5:337348.
- Beverly Trezek, Ye Wang, and Peter Paul. 2010. *Reading and deafness: Theory, research and practice*. Clifton Park, NY: Cengage Learning.
- Robert Van Valin and Randy LaPolla. 1997. *Syntax: Structure, Meaning, and Function*. Cambridge University Press.
- Schreuder R Knoors H Snik A. Vermeulen AM, van Bon W. 2007. Reading comprehension of deaf children with cochlear implants. *Journal of Deaf Studies and Deaf Education*, 12(3):283–302.
- Zhemina Zhu, Delphine Bernhard, and Iryna Gurevych. 2010. A monolingual tree-based translation model for sentence simplification. In *Proceedings of The 23rd International Conference on Computational Linguistics, Beijing, China*, pages 1353–1361.