

# Bidirectional Indonesian-Gorontalo Text Translator: Rule-Based Approach

*by* Wrastawa Ridwan

---

**Submission date:** 18-Apr-2023 03:53AM (UTC-0400)

**Submission ID:** 2068159437

**File name:** artikel\_IJAER.pdf (291.62K)

**Word count:** 2905

**Character count:** 16263

## 7 Bidirectional Indonesian-Gorontalo Text Translator: Rule-Based Approach

Wrastawa Ridwan\*, Rahmat D.R Dako

Department of Electrical Engineering

State University of Gorontalo

Jl. Jendral Sudirman no.6 Kota Gorontalo 96128, Indonesia

\*wridwan@ung.ac.id

**Abstract**—The objective of this research is to design the translation procedure from Indonesian language to Gorontalo language and vice versa. We then apply these procedures in computer programming algorithm to form bidirectional translator application software of Indonesian–Gorontalo language. The difference between the Indonesian and Gorontalo language is on the verbs that have several forms depending on the time in the active sentence. Besides, its form in present time is influenced by subject, whether it is singular or plural. Based on these differences, the design of algorithms for the translation method applied rule-based approach. The results of this study are the algorithms, text translator and its implementation in software. The test result shows that the application can generate a text translator translation sentence in accordance with the rules of the main grammar. Testing conducted on 4800 test sentences, correct sentences obtained 3603 (75.06%).

**Keywords** — rule-based, NLP, Indonesian, Gorontalo language

### I. INTRODUCTION

Natural language translators have been developed since the Georgetown-IBM system was demonstrated on 9<sup>th</sup> January 1954 [1]. De Silva et al. in [2] have developed a Sinhala to English language translator. Wijerathna et al. in [3] then developed a vice versa Sinhala-English translator (SEES). The SEES translator accuracy is 87%. The machine translation system, translated 500 grammatically well-structured Sinhala sentences to English and 150 grammatically well-structured English sentences to Sinhala. Furthermore, Soyusiawaty and Haspiyan in [4] developed a WAP-based (Wireless Application Protocol) software applications Indonesian to Sasak language dictionary.

In an effort to increase the success rate or accuracy in translation, some researchers applied method or approach to machine translation. Rule-based approach has applied in Arab<sup>3</sup> language translation [5,6]. Shaalan in [5] has addressed their successful efforts that involved rule-based approach for different Arabic natural language processing tasks. Shquier and Al Nabhan in [6], henceforth compiled test framework with rule-based approach for handling agreement (rules of grammar) and word-ordering (order of words) in the sentence translation from English to Arabic. The results were analyzed with the input sentence translation<sup>8</sup> certain words to put in some machine translation of which ALMUTARJIM Al Arabi, GOOGLE, TARJIM, SYSTRAN, and RBMT. From the analysis obtained RBMT (Rule Based Machine Translation)

produces a translation better than the other translation machine. The other researches of rule-based approach are in [2, 7, 9, 10, 11, 12]. In the same vein, Rhaman and Tarannum in [7] have developed a Bangla to English translator which is modified from their previous research in [8]. In [9], Centelles and Costa-jussa implemented the rule based approach for Chinese to Spanish machine translator, meanwhile Rajan et al. [10] applied such an approach in English to Malayalam machine translator.

Utami and Hartati in [11] found that the English to Indonesian text translator with rule-based method is able to translate the words in the "daily conversation" reasonably well with the corresponding translation of the sentence structure and approach the original meaning. On the other hand, the implementation of a rule based approach in English-Javanese was performed in [12]. The result of the research that the rule-based approach can handle differences of grammar of both languages so that produce the corresponding translation.

Based on the researches presented above, the rule based approach is also applied in this research. The translation using a rule based approach produces a translation based on language grammar. Application of the rule-based method was carried out the difference grammar between Indonesian and Gorontalo language. The fundamental differences of the two languages are the Indonesian language is not using form of time, meanwhile Gorontalo language uses it. The transformation of the form of time is marked by changes in affix of verbs. The change also affected by verb affixes singular and plural actors in active sentences. Other differences is the use of the article in the Indonesian language is not mandatory, but in Gorontalo language it should be used.

Dako et al. in [13] has conducted preliminary research to create software applications for Indonesian to Gorontalo language translator (one-way translator). This paper is a continuation of the research. The previous study result was one-way translator with success level of translation accuracy 71.85%. Further, this study develops two important things. First, to create the translation algorithms based on the procedures and rules of translation from Gorontalo language to Indonesian and adding some rules that have not been applied in previous studies. Second, to obtain the design software bidirectional Indonesian–Gorontalo language translator.

### II. PROBLEM FORMULATION

Based on the explanation above, we have formulated three problems of this research. They are first, Getting the steps or procedures in the Indonesian – Gorontalo language text translation vice versa; second, implementing steps or

procedures into an algorithm for developing a computer application software Indonesian – Gorontalo language translator; and finally, Getting a good translation results with rule-based method.

### III. DESIGN OF INDONESIAN – GORONTALO LANGUAGE TRANSLATION

The concept of translating sentences of Indonesian to Gorontalo language is performed by several stages:

- (1) The analysis and design phases of the system,
- (2) The implementation of translation rules-based on grammar rules,
- (3) Testing of the system by entering the input sentence in Indonesian and evaluate results of translation in Gorontalo and vice versa.

The translator system is designed following the steps below:

- 1) **Design of grouping words.** Grouping words in the dictionary in a designed system aimed to classify words into different specific markers of word classes. This is to facilitate the implementation of the rules of grammar in the translator application. Table 1 shows some of the words which are derived from the classes of words. Number 1 and 2 show as a single and a plural marker respectively. Table 1 is purposed for translation direction from Indonesian to Gorontalo's language. The grouping words of the other direction is presented in Table 2.
- 2) **Design of Table of words.** The table is designed for words database or dictionary. There are seven pieces of words table which is designed to accommodate the words database or dictionary.
- 3) **Analysis of Sentence Structure in Indonesian and Gorontalo language.** Translating sentences from Indonesian to Gorontalo language and vice versa can be done one to one or verbatim. To facilitate the process of synthesizing a sentence or phrase formed in Gorontalo language to Indonesian, it is necessary to compare the analysis of the composition of sentences or phrases of Indonesian and Gorontalo language.
- 4) **Design of translation algorithm.** Translation algorithm is separated into two parts of procedures based on the direction of translation. The first procedure is the Indonesian translation to Gorontalo language that could be described as follows:
  - a. Text input sentences parsed by a space character, then they are put in the dictionary to obtain equivalent words in Gorontalo language. Equivalent words are then reconstituted into a sentence.
  - b. The sentence translation, then examined sequentially through some procedures of the rule of Gorontalo language grammar which is presented in Fig. 1.

TABLE I. GROUPING TYPE OF WORDS IN INDONESIAN TO GORONTALO LANGUAGE TRANSLATION SYSTEM [13].

Type of word	Bookmarks	Information
Verb	(k)	Verb in active sentences
	(kp)	Verb in passive sentences
	(kk)	Verb form of the command.
Noun	(b)	Noun (concrete and abstract).
	(bm), (bm1), (bm2)	Noun related to people/humans.
	(bb), (bb1), (bb2)	Noun related with objects except humans which can serve as agents in a sentence.
	(bo), (bo1), (bo2)	Associated with the noun in kinship
Pronouns	(bw)	Noun related with time information
	(a)	general pronoun
Adverb	(ao1), (ao2)	Personal pronoun
	(e)	
Adjectives	(s), (s1), (s2)	
Numerals	(l), (l1), (l2)	
words Task	(t)	

TABLE II. GROUPING TYPE OF WORDS IN GORONTALO TO INDONESIAN TRANSLATION SYSTEM

Type of word	Bookmarks	Information
Verb	(k)	Verb in active sentences
	(ka)	Verb in passive sentences
Noun	(b)	General pronoun
Pronouns	(a)	
Adverb	(ao1), (ao2)	Personal pronoun
	(e)	
Adjective	(s)	
Numeral	(l)	
Word task	(t)	

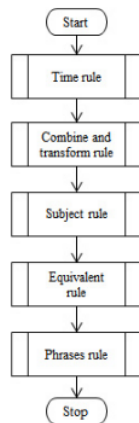


Figure 1. Flow chart that represents sequential procedures of rules for Indonesian to Gorontalo language translation.

General explanation of those procedures (see Fig. 1) will describe as follows:

- **Time rule.** Timing rules either past, present or future are the underlying for application to other rules. If the rules of the time have been addressed, it will be easier for handling of plural or singular verb changes. Pseudo code shown in Fig. 2.

For example, if an input sentence in Indonesian "Saya tiba di Jakarta kemarin". This sentence is translated one-by-one into Gorontalo "waatia medungga to Jakarta olaango". *Medungga* in Gorontalo is the present verb form. But, this sentence is followed by an adverb in the past *kemarin* = *olaango*, so *medungga* must be changed in the past form, by adding affixes *lo-* in front of the verb. It's become *lomedungga*. The word *lomedungga* will be changed to *ledungga* in the Equivalent Rule process.

- **Combine and transform rule.** This procedure will examine and incorporate the adverbs (*ma, he, lo*) which are followed by verbs, because in Gorontalo language these adverbs are the prefix of the verb. Pseudo code that describes this procedure in Fig. 3.

For example, the sentence in Indonesian *Kamar itu telah saya bersihkan* = *Huali boito lo popoberesio watia popoberesio*. This sentence is a passive sentence in Indonesian. However, there is a mistake in the translation in Gorontalo language. So it must be changed, by moving the word *watia* after the word *popoberesio*. It becomes *Huali boito lo popoberesio watia* and the bookmark of active verbs changed to the passive verbs.

- **Subject rule.** The procedure will handle words related to subject as an actor in a sentence. For example, word *tau* which mean *orang* in Indonesian has several forms depends on the words that are following it. If it is followed by word *boito* or *botia* then it must change to *ta*. Beside that this procedure is also handling plural or singular form of the noun as a subject in a sentence. The General pseudo code for this procedure can be seen in Fig. 4.

```

DO UNTIL LAST WORD
  EXAMINE word(i) = past bookmark
  IF bookmark word(i) = past time THEN
    EXAMINE CONDITION
    IF word(i) = verb AND
       word(i-1) <> "lo" THEN
      ADD kata "lo" + verb
    ELSE
      EXIT PROCEDURE
    ENDIF
  ENDIF
RESTRUCTURE Gorontalo Sentence
    
```

Figure 2. Procedure of time rule.

```

DO UNTIL LAST WORD
  EXAMINE word(i) and word(i+1)
  IF word(i) = adverb THEN
    IF word(i+1) = noun(subject) THEN
      IF word(i+2) = verb THEN
        POSITION EXCHANGE word(i+1) and
          word(i+2)
        CHANGE bookmark word(i) (active
          verb = passive verb)
        ELSE IF kata(i+1) = verb THEN
          COMBINE word(i) and word(i+1)
        ENDIF
      ENDIF
    ENDIF
  ENDIF
    
```

Figure 3. Procedure of combine and transform rule.

- **Equivalent rule.** In this procedure, the subject as an actor is specified in the active sentence both in plural or singular. In passive sentences, this process is not implemented, because it does not affect the verb. The process will continue to search the equivalent verb in words table or dictionary. This searching is intended to find the equivalent word, because in the previous process, there is a presence of the plural and singular actor and the combining of adverbs and verbs. Pseudo code of this procedure in Fig 5.

For example, *sedang makan* = *hemonga*. If the verbal phrase is preceded by plural actors such as in Indonesian *mereka* = *timongolio* in Gorontalo, then *timongolio hemonga* must be changed into *timongolio hipongala*.

```

DO UNTIL LAST WORD
EXAMINE word(i)
IF word(i) = noun(human(name, kinship)) _
    THEN
    IF word(i-1) <> "wole" or "woli" or _
        "wolemei" or "wolilei" or _
        "te" or "ti"
        ADD "te" or "ti" in front of _
        word(i)
    ENDIF
    IF (word(i) = "tau-tauwaalo" or _
        "tau") AND (kata(i+1)="boito" or _
        "botia") THEN
        CHANGE word(i) = "ta"
    ENDIF
    IF kata(i) = noun(subject) THEN
        EXAMINE bookmark word(i) = plural _
        or singular
        IF word(i+1) = verb THEN
            MARK word(i+1)
            MARK word(i) as "singular" or _
            "plural"
        ENDIF
    ENDIF
    FIND equivalent word(i) = verb in _
    dictionary
    RESTRUCTURE Gorontalo sentence
    
```

Figure 4. Procedure of subject rule.

```

DO UNTIL LAST WORD
EXAMINE word(i)
IF word(i) = noun(human(name, kinship)) _
    and word(i) = personal pronoun THEN
    IF bookmark word(i) = "plural" THEN
        FIND equivalent word IN plural dictionary table
        ELSEIF bookmark word (i) = "singular" _
        THEN
            FIND equivalent word IN singular _
            dictionary table
        ENDIF
    ENDIF
    FIND equivalent word(i) = verb IN _
    dictionary
    RESTRUCTURE Gorontalo sentence
    
```

Figure 5. Procedure of equivalent rule.

- **Phrases rule.** Phrases procedure will handle some rules of phrases. Pseudo code for this procedure in Fig. 6.

Furthermore, translation algorithms for direction from Gorontalo language to Indonesian are:

- The input sentences parsed by a space character, then the words are put in the dictionary to obtain its equivalent words in the Indonesian language, and the words then restructure into Indonesian sentence.
- This translation sentence also examined sequentially through three sequences rules procedures in Indonesian grammar as shows in Fig. 7.

```

DO UNTIL LAST WORD
EXAMINE word(i)
If word(i) = noun THEN
    IF word(i+1) = personal pronoun THEN
        JOIN word(i) + word(i+1)
        ELSEIF word(i+1) = noun
            PUT "li" or "lo" between words
        ENDIF
    ELSEIF word(i) = verb(active) THEN
        IF word(i+1) = personal pronoun THEN
            CHANGE personal pronoun = object form
        ELSEIF word(i+1) = noun(human(name, kinship)) THEN
            PUT "li" between words
        ENDIF
    ELSEIF word(i) = verb(passive) THEN
        IF word(i+1) = personal pronoun THEN
            CHANGE personal pronoun = objectform
            JOIN word(i) = word(i)+ word(i+1)
        ELSEIF word(i+1) = noun(human(name, kinship)) THEN
            PUT "li" between words
        ELSEIF word(i+1) = noun concrete/abstract)
            PUT "lo" between words
        ENDIF
    ELSEIF word(i) = "to" THEN
        IF word(i+1) = "teea" THEN
            CHANGE word(i+1) = "teeamai"
        ELSEIF word(i+1) = "teeto" THEN
            CHANGE word(i+1) = "teeto(ma'o/mota/mola)"
        ENDIF
    ELSEIF word(i) = numeral THEN
        IF word(i+1) = "hungo" or word(i+1) = "patahu" THEN
            EXCHANGE POSITION word(i+1) and word(i)
            word(i) = "mehelu"
        ENDIF
    ENDIF
    RESTRUCTURE Gorontalo sentence
    
```

Figure 6. Procedure of phrases rule.

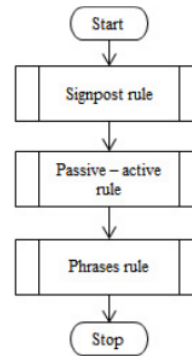


Figure 7. Flow chart that represents sequential procedures of rules for gorontalo language to Indonesian translation.

General description of the procedures in Fig. 7 is described as follows:

- **Signpost rule.** In Gorontalo language, signpost has a lot of functions. Signpost can serve as a word if it is positioned before the affirmed words and serves as a morpheme after the affirmed words. This procedure examines the signpost position either before or after its affirmed word and transforms it according to Indonesian grammar. General Pseudo code for this procedure in Fig. 8.
- **Passive – active rule.** Design of passive-active procedure is intended to handle some of form changes of the basic verbs that can be placed as an active sentence, and could also be used in passive sentences. Pseudo code for this procedure in Fig. 9.

- **Phrases rule.** Procedures for handling phrases in Gorontalo language into Indonesian is similar to the translation procedures phrases that have been designed on the rules of procedure phrase translation from Indonesian into Gorontalo language. This procedure is made separately in order to simplify when writing program codes for phrases that have not been identified at this stage of the designed system. Some of the rules that have been handled are described in pseudo code in Fig. 10.

```
DO UNTIL LAST WORD
EXAMINE word(i)
IF word(i) = "datang(mpa)" or _
    word(i) = "pergi(mpa)" THEN
    IF right(word(i + 1), 3) = "(k)" and
        word(i) = "pergi(k)" or "datang(k)" THEN
        word(i) = ""
    ENDIF
ENDIF
```

Figure 8. Procedure of signpost rule.

```
DO UNTIL LAST WORD
EXAMINE word(i)
IF right(word(i),4) = "(ka)" THEN
    IF word(i + 1) = "le" or "li" or "lo" or _
        "lemei" or "lilei" THEN
        word(i) = "di" + kata(i) + "(kp)"
    ENDIF
ENDIF
IF right(word(i),4) = "(kp)" THEN
    IF word(i + 1) = "le" or "li" or "lo" or _
        "lemei" or "lilei" THEN
        word(i + 1) = "oleh/dengan"
    ENDIF
ENDIF
IF right(word(i),3) = "(k)" THEN
    IF word(i + 1) = "ole" or "oli" or
        "olemei" or "olilei" THEN
        word(i + 1) = ""
    ENDIF
ENDIF
RESTRUCTURE Gorontalo sentence
```

Figure 9. Procedure of passive-active rule

```
DO UNTIL LAST WORD
EXAMINE word(i)
Select case word(i)
Case is = "di(t)"
    IF word(i+1) = "le" or "li" THEN
        word(i) = "pada/sama"
    ENDIF
Case is = "ke(t)"
    IF word(i+1) = "le" or "li" THEN
        word(i) = "kepada/untuk"
    ENDIF
Case is = "dari(t)"
    IF word(i+1) = "ole" or "oli" THEN
        word(i+1) = ""
    ENDIF
Case is = "mehelu"
    IF right(word(i+1)) = "(l)" THEN
        word(i) = word(i+1)
        word(i+1) = "buah/ekor"
    ENDIF
Case is = "sini" or "sana"
    IF word(i+1) = "le" or "li" or "lo" THEN
        word(i+1) = "dari"
    ENDIF
End select
IF word(i+1) = "ole" or "oli" THEN
    IF right(word(i)) = "(k)" THEN
        word(i+1) = ""
    ElseIf right(word(i)) = "(ka)" THEN
        word(i+1) = "oleh"
    ENDIF
IF word(i+1) = "lo" THEN
    IF right(word(i)) = "(k)" or "(ka)" THEN
        word(i+1) = "dengan/oleh"
    ENDIF
ENDIF
ENDIF
RESTRUCTURE Gorontalo sentence
```

Figure 10. Procedure of phrases rule.

#### IV. RESULT

To measure the level of success or accuracy of the translation, the application is tested by inserting sentences of Indonesian or Gorontalo language as input to the application. Then the output sentences of translation in Indonesian and Gorontalo language are analyzed.

The tests were carried out in two steps. The first step was conducted by asking 80 respondents in 4 different districts. Each respondent was asked to enter a total of 60 sentences, 30 sentences for each Indonesian and Gorontalo language. The total sentence obtained in this first phase of testing was 4800 sentences. There are 2241 sentences which are correct and the translation accuracy is 46.69%. Based on the result, the translation system, then modified adding vocabulary to the dictionary and enhanced some rules that have not been applied to the system. The second step of testing was done by inserting back those sentences. The results, the truth is obtained by an increase in the level of success or accuracy of translation, which is 75.06%. The translation faults are due to many words that have not been inserted into the dictionary, some words depend on the context of the sentence, and some translation rules which have not been implemented because it has not been identified during the design process.

#### V. CONCLUSION

In this research, we have identified the translation steps of Indonesian and Gorontalo language based on grammar rules in the form of algorithm translator. Then, the algorithm translator algorithm is implemented in computer applications with the best accuracy of the translation is 75.06%.

#### ACKNOWLEDGMENT

The authors are grateful to the Research Institute of State University of Gorontalo and Directorate General of Higher Education – Ministry of Education and Culture Republic of Indonesia for providing financial support to complete this research.

#### REFERENCES

- [1] Hutchins, J., 2004, "The Georgetown-IBM experiment demonstrated in January 1954". 6th Conference of the Association for Machine Translation in the Americas, AMTA, pp.102-114.
- [2] De Silva, D., Alahakoon, A., Udayangani, I., Kumara, V., Kolonnage, D., Perera, H., and Thelijjagoda, S., 2008, "Sinhala to english language translator". 4th Int. Conf. Information and Automation for Sustainability (ICIAFS), pp.419-424.
- [3] Wijerathna, L., Somaweera, W.L.S.L., Kaduruwana, S.L., Wijeshinghe, Y.V., De Silva, D.I., Pulasinghe, K., and Thelijjagoda, S., 2008, "A Translator from Sinhala to English and English to Sinhala (SEES)". Int. Conf. Tech. for Emerging Regions (ICTer), pp.14-18.
- [4] Soyusiawaty, D., and Haspiyan, R., 2009, "Aplikasi kamus bahasa indonesia – bahasa sasak berbasis WAP". Seminar Nasional Informatika 2009 (semnasIF 2009).
- [5] Shaalan, K., 2010, "Rule-Based Approach in Arabic Natural Language Processing". Int. J. on Information and Communication Tech., 3(3).
- [6] Shquier, A., and Nabhan, A.L., 2010, "Rule-based approach to tackle agreement and word-ordering in english-arabic machine translation", European & Mediterranean Conf. Informatics Sys. (EMCIS).
- [7] Rhaman, M.K., and Tarannum, M., 2012, "A Rule Based Approach for implementation of Bangla to English Translation". Int. Conf. Adv. Computer Sci. Appl. and Tech (ACSAT), pp. 13-18.
- [8] Rhaman, M.K., and Tarannum, M., 2010, "An initiative bangla english natural language translation using case". Int. Conf. Audio Language and Image Processing (ICALIP), pp. 1106-1111.
- [9] Centelles, J. and Costa-jussa, M.R., 2014, "Chinese-to-Spanish rule-based machine translation system", The 3<sup>rd</sup> Workshop on Hybrid Approaches to Translation (HyTra), pp. 78-82.
- [10] Rajan, R., Sivan, R., Ravindran, R., and Soman, K.P., 2009 "Rule Based Machine Translation from English to Malayalam". Int. Conf. Adv. in Comp., Control & Telecommunication Tech. (ACT), pp. 439-441.
- [11] Utami, E., and Hartati, S., 2007, "Pendekatan metode rule based dalam mengalihbahasakan teks bahasa inggris ke teks bahasa indonesia". Informatics Journal, 8(1), pp. 42-53.
- [12] Wikantyasning, N., 2005, "Penerjemah inggris – jawa bagi siswa asing menggunakan metode rule based", Master Thesis, Teknik Elektro. Universitas Gadjah Mada.
- [13] Dako, R.D.R., Ridwan, W., and Dako, R. T., 2013, "Indonesian to gorontalo text translator". Int. J. Research and Rev. in App. Sci, 16( 2), pp. 219-223.

# Bidirectional Indonesian-Gorontalo Text Translator: Rule-Based Approach

## ORIGINALITY REPORT

4%

SIMILARITY INDEX

4%

INTERNET SOURCES

3%

PUBLICATIONS

1%

STUDENT PAPERS

## PRIMARY SOURCES

1

[www.icter.org](http://www.icter.org)

Internet Source

1%

2

[biblio.ugent.be](http://biblio.ugent.be)

Internet Source

1%

3

[www.semanticscholar.org](http://www.semanticscholar.org)

Internet Source

1%

4

[www.iiste.org](http://www.iiste.org)

Internet Source

<1%

5

M I Bahua, D O Suparwata. "The level of society's participation in critical land rehabilitation in DAS Randangan Pohuwato Regency", IOP Conference Series: Earth and Environmental Science, 2018

Publication

<1%

6

[www.iseing.org](http://www.iseing.org)

Internet Source

<1%

7

[www.ung.ac.id](http://www.ung.ac.id)

Internet Source

<1%



8

Mohammed M. Abu Shquier, Tengku Mohd T. Sembok. "Word agreement and ordering in English-Arabic machine translation", 2008 International Symposium on Information Technology, 2008

Publication

<1 %

9

L. Wijerathna, W.L.S.L. Somaweera, S.L. Kaduruwana, Y.V. Wijesinghe, D.I. De Silva, K. Pulasinghe, S. Thellijagoda. "A Translator from Sinhala to English and English to Sinhala (SEES)", International Conference on Advances in ICT for Emerging Regions (ICTer2012), 2012

Publication

<1 %

10

[computerresearch.org](http://computerresearch.org)

Internet Source

<1 %

11

[media.neliti.com](http://media.neliti.com)

Internet Source

<1 %

Exclude quotes On

Exclude matches Off

Exclude bibliography On