



Implementasi Metode RBMT dalam Penerjemahan Bahasa Indonesia ke Bahasa Makassar

Wan Muhammad Hanif, Yusra*, Muhammad Fikry, Febi Yanto, Siska Kurnia Gusti

Fakultas Sains dan Teknologi, Program Studi Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: ¹12150113633@students.uin-suska.ac.id, ^{2,*}yusra@uin-suska.ac.id, ³muhammad.fikry@uin-suska.ac.id,

⁴febiyanto@uin-suska.ac.id, ⁵siskakurniagusti@uin-suska.ac.id

Email Penulis Korespondensi: yusra@uin-suska.ac.id

Abstrak—Penelitian ini dilakukan untuk mengatasi keterbatasan sumber daya bahasa daerah, khususnya Bahasa Makassar, yang belum memiliki penerjemahan otomatis yang memadai. Permasalahan dalam penelitian ini adalah belum tersedianya terjemahan otomatis yang memadai untuk Bahasa Makassar. Penelitian ini bertujuan untuk menerapkan metode penerjemahan berbasis aturan dalam menerjemahkan teks dari Bahasa Indonesia ke Bahasa Makassar. Penelitian ini berfokus pada implementasi metode *Rule Based Machine Translation* (RBMT) untuk menerjemahkan teks dari Bahasa Indonesia ke Bahasa Makassar dengan memanfaatkan bahasa pemrograman Python. Implementasi RBMT dilakukan melalui proses tokenisasi, analisis morfologi, pencocokan kosakata, serta penerapan aturan gramatis yang mencakup identifikasi awalan dan akhiran. Data yang digunakan terdiri atas kamus dwibahasa yang disusun dari berbagai sumber serta kumpulan kalimat uji yang mewakili struktur kalimat sehari-hari. Untuk pengujian terjemahan menggunakan metode *Word Error Rate* (WER) dengan hasil 0.289 dan menggunakan metode *Character Error Rate* (CER) dengan hasil pengujian sebesar 0.21, yang berdasarkan rentang penilaian termasuk dalam kategori “Baik”. Hasil utama penelitian menunjukkan bahwa penerapan metode RBMT mampu menghasilkan terjemahan yang cukup akurat pada tingkat kata dan karakter. Temuan tersebut memperlihatkan bahwa pendekatan berbasis aturan dapat diterapkan secara efektif pada bahasa daerah yang memiliki keterbatasan data digital dan memberikan gambaran awal pemanfaatan metode berbasis aturan untuk mendukung pengembangan dan pelestarian bahasa daerah.

Kata Kunci: Penerjemahan Berbasis Aturan; Bahasa Makassar; Morfologi; Kosakata; Penerjemahan Otomatis

Abstract—This research was conducted to address the limited availability of linguistic resources for regional languages, particularly Makassar Language, which does not yet have adequate automatic translation support. The main problem addressed in this study is the absence of a reliable automatic translation system for Makassar Language. The objective of this research is to apply a rule-based translation method to translate text from Indonesian into Makassar Language. This study focuses on the implementation of the Rule-Based Machine Translation (RBMT) method for translating Indonesian text into Makassar Language using the Python programming language. The RBMT implementation involves tokenization, morphological analysis, vocabulary matching, and the application of grammatical rules, including the identification of prefixes and suffixes. The data used consist of a bilingual dictionary compiled from various sources and a set of test sentences representing everyday sentence structures. Translation evaluation was carried out using the Word Error Rate (WER) method, yielding a result of 0.289, and the Character Error Rate (CER) method, with a result of 0.21, which fall into the “Good” category based on the evaluation scale. The main findings indicate that the application of the RBMT method is capable of producing reasonably accurate translations at both the word and character levels. These findings demonstrate that a rule-based approach can be effectively applied to regional languages with limited digital data and provide an initial overview of the potential use of rule-based methods to support the development and preservation of regional languages.

Keywords: Rule Based Translation; Makassarese Language; Morphology; Lexical Analysis; Automatic Translation

1. PENDAHULUAN

Penerjemahan merupakan proses kompleks yang tidak hanya mengalihkan kata dari bahasa sasaran, tetapi juga mempertahankan kesepadan makna, struktur sintaksis, serta konteks budaya yang terkandung dalam teks sumber. Sebagaimana dinyatakan oleh [1] penerjemahan menuntut pemindahan pesan secara akurat dari frasa, klausu, hingga paragraf, sehingga makna utuh teks dapat diterima secara serasi oleh pembaca bahasa sasaran. Proses ini tidak bersifat mekanis, karena perbedaan sistem bahasa dan budaya berpotensi menimbulkan pergeseran makna berupa pengurangan, penambahan, maupun perubahan informasi dalam hasil terjemahan [2]. Oleh karena itu, peran penerjemah tidak berhenti pada aspek linguistik, melainkan juga menyesuaikan struktur dan pilihan leksikal bahasa sasaran agar pesan dalam teks sumber dapat dikomunikasikan secara alami dan dipahami oleh pembaca sasaran [3].

Di Indonesia, tantangan penerjemahan semakin besar ketika berkaitan dengan bahasa daerah. Indonesia memiliki lebih dari 700 bahasa lokal yang mencerminkan keragaman linguistik dan budaya. Namun, penggunaan bahasa daerah saat ini semakin menurun dan mulai terlupakan [4]. Perkembangan teknologi komunikasi dan media sosial turut memperkuat fenomena ini, karena penggunaan bahasa lebih banyak terpusat pada bahasa nasional dan global. Oleh sebab itu, upaya digitalisasi dan teknologi pelestarian bahasa daerah menjadi semakin penting untuk mencegah pergeseran yang lebih jauh.

Salah satu bahasa daerah yang memiliki kompleksitas linguistik dan menghadapi tantangan pelestarian adalah Bahasa Makassar. Bahasa ini memiliki karakteristik morfologi yang kaya, melibatkan berbagai jenis afiksasi, serta proses fonologi khas seperti *nasalization* [5]. Dalam kajian linguistik, morfologi dipahami sebagai cabang ilmu yang mempelajari pembentukan kata melalui interaksi morfem, termasuk afiksasi, reduplikasi, dan perubahan morfonemik, yang berperan penting dalam menentukan makna dan fungsi gramatiskal suatu kata [6]. Hasil penelitian perbandingan bahasa menunjukkan adanya perbedaan bentuk imbuhan antara Bahasa Indonesia dan Bahasa Makassar [7]. Perbedaan



ini bukan hanya bersifat bentuk, tetapi juga melibatkan fungsi yang maknanya tidak selalu sama. Kompleksitas tersebut membuat proses penerjemahan memerlukan penanganan linguistik yang lebih dalam, terutama ketika menggunakan pendekatan berbasis data besar yang memerlukan korpus paralel dalam jumlah banyak.

Namun, ketersediaan sumber daya linguistik Bahasa Makassar masih sangat terbatas. Minimnya korpus paralel, kamus digital, dan sumber data terstruktur menyebabkan teknologi penerjemahan berbasis Neural Machine Translation (NMT) atau Statistical Machine Translation (SMT) sulit diterapkan secara optimal, pendekatan berbasis data tersebut secara umum membutuhkan korpus paralel berskala besar agar mampu mempelajari pola terjemahan secara akurat, sehingga bahasa dengan sumber daya terbatas cenderung menghasilkan kualitas terjemahan yang rendah [8]. Meskipun NMT dan Large Language Models (LLM) saat ini berkembang pesat dan banyak digunakan, efektivitas metode tersebut sangat bergantung pada ketersediaan data dalam jumlah besar, yang pada Bahasa Makassar masih belum tersedia. Oleh karena itu, pemilihan Rule Based Machine Translation (RBMT) dalam penelitian ini didasarkan pada pertimbangan keterbatasan sumber daya kebahasaan, karena RBMT tidak memerlukan kumpulan data terjemahan berskala besar dan lebih sesuai untuk bahasa daerah yang memiliki keterbatasan data serta karakteristik morfologi yang kompleks. Penelitian penerjemahan statistik pada bahasa daerah juga menunjukkan bahwa performa SMT sangat bergantung pada jumlah korpus dan masih menghasilkan akurasi yang relatif moderat berdasarkan metrik BLEU [9]. Sebaliknya, pendekatan RBMT terbukti mampu menghasilkan terjemahan yang sesuai dengan kaidah linguistik lokal meskipun tanpa dukungan korpus paralel berskala besar, sebagaimana ditunjukkan pada penerjemahan Bahasa Indonesia ke Bahasa Ende Lio [10]. Pendekatan berbasis aturan ini juga efektif dalam konteks pelestarian bahasa dan aksara daerah, seperti pada penelitian transliterasi aksara Bali berbasis Rule Based yang dapat beroperasi secara offline tanpa ketergantungan pada korpus linguistik berskala besar [11]. Kesenjangan antara kebutuhan pelestarian bahasa daerah dan dukungan teknologi modern menjadi isu utama dalam pengembangan teknologi kebahasaan di Indonesia.

Untuk bahasa yang memiliki keterbatasan sumber daya digital dan data kebahasaan, pendekatan RBMT menjadi relevan karena mengandalkan aturan linguistik eksplisit dan kamus dwibahasa. RBMT terdiri dari serangkaian aturan morfologis, sintaksis, dan leksikal yang dirumuskan oleh ahli bahasa, sehingga tidak bergantung pada data pelatihan yang besar [12], [13]. Secara umum, RBMT dikembangkan melalui tiga pendekatan utama, yaitu direct translation, transfer based translation, dan interlingua based translation. Pendekatan direct translation menerjemahkan teks berdasarkan kamus dengan analisis minimal sehingga struktur bahasa sumber banyak dipertahankan, namun hasil terjemahan cenderung kurang alami [14]. Pendekatan interlingua menggunakan representasi abstrak antarbahasa yang efektif untuk terjemahan multibahasa, tetapi sulit diterapkan pada bahasa dengan struktur morfologis yang kompleks. Oleh karena itu, pada bahasa daerah yang memiliki morfologi kaya dan afiksasi kompleks seperti Bahasa Makassar, pendekatan transfer based translation lebih sesuai karena mampu mengakomodasi transformasi afiksasi dan struktur sintaksis secara sistematis.

Penelitian terdahulu memperlihatkan keberhasilan RBMT pada bahasa daerah maupun bahasa minoritas. Implementasi RBMT untuk Bahasa Aceh menggunakan aturan perubahan suku kata untuk menangani afiksasi dan variasi bunyi dalam proses penerjemahan [15]. Penelitian pada dialek Tunisia juga memperlihatkan skor *Bilingual Evaluation Understudy* (BLEU) yang baik setelah menerapkan lebih dari 500 aturan struktural [16]. Selain penerjemahan, RBMT sukses dalam transliterasi aksara Jawa ke huruf Latin dengan akurasi mencapai 96,44% [4]. Bahkan, pendekatan *hybrid* yang menggabungkan RBMT dan *Large Language Model* (LLM) seperti yang dilakukan oleh [12] menunjukkan kemajuan signifikan dalam revitalisasi bahasa bahasa yang memiliki keterbatasan sumber daya digital dan data kebahasaan. Penerapan RBMT pada bahasa daerah Indonesia juga menunjukkan hasil yang baik, penelitian pada penerjemahan Bahasa Indonesia ke Bahasa Minangkabau mencapai akurasi tinggi berdasarkan metode Word Error Rate (WER) [17], sementara implementasi RBMT pada Bahasa Alas Aceh Tenggara membuktikan bahwa pendekatan berbasis aturan efektif digunakan pada bahasa dengan keterbatasan sumber daya linguistik [18]. Di sisi lain, implementasi penerjemah bahasa daerah berbasis web seperti pada Bahasa Tetun Indonesia telah dilakukan dengan memanfaatkan kamus dwibahasa untuk mendukung pembelajaran kosakata, namun pendekatan tersebut belum menerapkan analisis morfologis berbasis aturan maupun evaluasi akurasi terjemahan secara kuantitatif [19]. Secara keseluruhan, studi-studi tersebut memberi bukti bahwa RBMT merupakan solusi efektif untuk bahasa yang tidak memiliki data besar.

Meskipun banyak keberhasilan RBMT dilaporkan, penelitian yang menerapkan RBMT pada Bahasa Makassar masih sangat terbatas. Padahal, Bahasa Makassar memiliki dokumentasi kamus dasar yang dapat menjadi dasar Implementasi RBMT. Misalnya, kamus Indonesia–Makassar yang disusun oleh Balai Penelitian Bahasa menyediakan kosakata dasar yang diperlukan untuk membangun komponen leksikon dalam RBMT. Selain itu, analisis perbandingan bahasa mengenai perubahan afiksasi [7] memberikan landasan kuat untuk merumuskan aturan transformasi morfologis yang presisi. Ketersediaan bahan linguistik dasar inilah yang sebenarnya memungkinkan penerapan RBMT pada Bahasa Makassar, meskipun korpus paralel tidak tersedia.

Kesenjangan atau gap penelitian muncul pada tiga aspek. Pertama, belum ada penelitian yang secara sistematis mengimplementasikan metode RBMT untuk penerjemahan otomatis dari Bahasa Indonesia ke Bahasa Makassar, padahal bahasa sasaran memiliki sistem morfologi yang sangat kaya sehingga ideal untuk pendekatan berbasis aturan. Kedua, belum ada evaluasi kuantitatif berbasis metrik *Word Error Rate* (WER) atau *Character Error Rate* (CER) untuk mengukur akurasi penerjemahan Bahasa Makassar, sementara metrik tersebut telah banyak digunakan dalam evaluasi kinerja sistem penerjemahan dan pengolahan bahasa lainnya [20][21]. Ketiga, sebagian besar penelitian sebelumnya fokus pada penerjemahan dua arah atau transliterasi, sementara kebutuhan pelestarian bahasa justru menuntut penerjemahan satu arah dari Bahasa Indonesia ke Makassar untuk mendukung pembelajaran bahasa daerah oleh penutur muda.



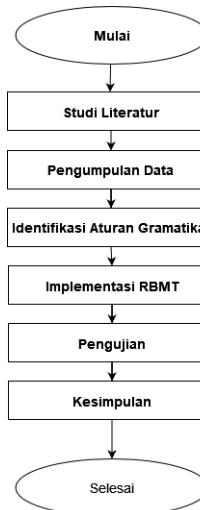
Penelitian ini hadir untuk mengisi celah tersebut dengan mengimplementasikan terjemahan berbasis aturan dari Bahasa Indonesia ke Bahasa Makassar. Implementasi terjemahan menggunakan aturan morfologi yang dirumuskan berdasarkan kamus Bahasa Makassar serta temuan linguistik mengenai proses afiksasi. Selain itu, penelitian ini mengevaluasi akurasi terjemahan dengan menggunakan metrik WER dan CER yang memungkinkan pengukuran kesalahan baik pada level kata maupun karakter. Kontribusi utama penelitian ini mencakup: (1) perumusan aturan RBMT yang sesuai dengan karakteristik afiksasi Bahasa Makassar; (2) implementasi metode RBMT untuk penerjemahan Bahasa Indonesia ke Bahasa Makassar tanpa memanfaatkan korpus paralel; (3) evaluasi hasil penerjemahan menggunakan metrik *Word Error Rate* (WER) dan *Character Error Rate* (CER) untuk memberikan gambaran objektif mengenai kualitas terjemahan yang dihasilkan.

Dengan demikian, penelitian ini tidak hanya berkontribusi pada pengembangan teknologi penerjemahan bahasa daerah, tetapi juga memiliki signifikansi sosial dalam mendukung pelestarian Bahasa Makassar melalui pendekatan komputasional yang terukur dan dapat direplikasi untuk bahasa-bahasa daerah lainnya.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Tahapan penelitian merupakan rangkaian langkah-langkah yang dilakukan secara sistematis untuk mencapai tujuan penelitian. Setiap tahapan dirancang untuk mendukung proses penelitian yang efektif dan memastikan bahwa hasil yang diperoleh valid dan dapat di pertanggung jawabkan. Adapun tahapan-tahapan penelitian dalam penelitian ini dapat dilihat pada Gambar 1.



Gambar 1. Alur Tahapan Penelitian

2.2 Studi Literatur

Studi literatur ini mengkaji berbagai penelitian terkait penerjemahan menggunakan metode RBMT. khususnya dalam konteks penerjemahan antara Bahasa Indonesia dan Bahasa Makassar. Penelitian-penelitian sebelumnya menunjukkan bahwa penerjemahan berbasis aturan gramatikal sangat bergantung pada pemahaman tentang morfologi kedua bahasa, terutama pada aspek awalan dan akhiran. Morfologi ini memainkan peran penting dalam membentuk makna dan struktur kata dalam kedua bahasa. Studi-studi ini memberikan dasar teori yang mendalam untuk pengembangan sistem RBMT dalam menerjemahkan Bahasa Indonesia ke Bahasa Makassar.

2.3 Pengumpulan Data

Data yang digunakan dalam penelitian ini terdiri dari data kosakata dan data uji kalimat. Data kosakata Bahasa Indonesia Bahasa Makassar diperoleh dari penelitian sebelumnya sebanyak 3.655 kata, kemudian diproses ulang untuk memperbaiki ketidak teraturan penulisan. Hasil pengolahan digunakan sebagai kamus penerjemahan dengan total 4.021 kata.

Data uji berupa 50 kalimat yang umum digunakan dalam percakapan sehari-hari. Data ini diperoleh dari buku *Tata Bahasa Makassar* terbitan Departemen Pendidikan dan Kebudayaan serta *Percakapan Bahasa Makassar* karya Barbara Friberg dan Kembong Daeng. Pemilihan data uji disesuaikan dengan karakteristik metode RBMT yang menerjemahkan teks secara per kata berdasarkan aturan dan kosakata yang tersedia.

2.4 Identifikasi Aturan Gramatikal

Pada metode RBMT, aturan gramatikal yang diterapkan dalam penelitian ini berfokus pada perubahan awalan dan akhiran dari Bahasa Indonesia ke Bahasa Makassar. Aturan tersebut disusun berdasarkan referensi Kamus Makasar Indonesia dan digunakan sebagai dasar dalam proses penerjemahan. Aturan gramatikal mencakup perubahan bentuk awalan dan akhiran



agar hasil terjemahan sesuai dengan struktur morfologi Bahasa Makassar. Ringkasan aturan perubahan awalan dan akhiran digunakan sebagai acuan dalam implementasi RBMT.

2.4.1 Morfologi Awalan

Perubahan awalan dari Bahasa Indonesia ke Bahasa Makassar ditentukan oleh huruf awal kata dasar. Awalan seperti me-, mem-, men-, dan ber- berubah menjadi bentuk awalan yang diawali huruf vokal. Jika kata dasar diawali huruf p, t, c, atau s, maka huruf awal tersebut digandakan, misalnya kata pate berubah menjadi appate. Jika kata dasar tidak diawali huruf tersebut, maka tidak terjadi penggandaan. Awalan lain seperti di-, ke-, dan se- berubah menjadi bentuk tetap dalam Bahasa Makassar tanpa perubahan tambahan. Ringkasan perubahan awalan ditampilkan pada Tabel 1.

Tabel 1. Perubahan Awalan

Nomor	Awalan	Perubahan
1	Me	aK-
2	Mem	aK-
3	Men	aK-
4	Meng	aN
5	Ber	aK-
6	Pe	pa
7	Pem	paK
8	Pen	paN
9	Peng	paN
10	Ter-	taK
11	Di-	Ri
12	Se-	si-
13	Ke-	ka-

2.4.2 Morfologi Akhiran

Selain awalan, Bahasa Makassar juga memiliki aturan perubahan akhiran. Beberapa akhiran Bahasa Indonesia seperti -nya, -lah, -kah, -mu, -kan, dan -an mengalami perubahan bentuk agar sesuai dengan kaidah Bahasa Makassar. Contohnya, akhiran -nya berubah menjadi -na, -mu menjadi -nu, serta -kan dan -an berubah menjadi -ang. Perubahan akhiran dilakukan secara langsung dan tidak dipengaruhi oleh huruf awal kata dasar. Ringkasan perubahan akhiran ditampilkan pada Tabel 2.

Tabel 2. Perubahan Akhiran

Nomor	Akhiran	Perubahan
1	Nya	Na
2	Lah	Mi
3	Kah	Ji
4	Mu	Nu
5	Kan	Ang
6	I	I
7	Ku	Ku

2.5 Implementasi RBMT

Implementasi metode RBMT pada penelitian ini dilakukan dengan menerapkan tahapan-tahapan penerjemahan secara sistematis berdasarkan aturan linguistik dan kamus yang telah disiapkan. Proses penerjemahan dimulai dari penyediaan data kamus sebagai sumber utama pencocokan kosakata, kemudian dilanjutkan dengan pemrosesan teks masukan melalui tokenisasi, analisis morfologi, serta penerapan aturan gramatiskal. Seluruh tahapan tersebut dirancang untuk menghasilkan terjemahan dari Bahasa Indonesia ke Bahasa Makassar yang sesuai dengan kaidah morfologi bahasa sasaran dan dapat diproses secara otomatis.

2.5.1 Pengambilan Data Kamus

Tahap awal implementasi RBMT adalah pemuatan kamus Bahasa Indonesia–Bahasa Makassar. Kamus disimpan dalam bentuk berkas CSV yang memuat kosa kata bahasa Indonesia dan Bahasa Makassar.

2.5.2 Tokenisasi Teks

Setelah kamus dimuat, proses penerjemahan menerima input berupa kalimat Bahasa Indonesia. Kalimat tersebut kemudian diproses melalui tahap tokenisasi untuk memisahkan teks menjadi unit-unit kata dan tanda baca.

2.5.3 Analisis Morfologi Kata

Analisis morfologi merupakan tahapan utama dalam implementasi RBMT yang bertujuan untuk mengenali struktur



morfologis kata dalam Bahasa Indonesia sebelum dilakukan proses penerjemahan. Tahapan ini difokuskan pada identifikasi kata turunan yang mengandung awalan, akhiran, atau kombinasi keduanya, berdasarkan aturan gramatikal yang telah diidentifikasi.

2.5.4 Penerapan Aturan Gramatikal

Setelah kata dasar dan awalan akhiran dikenali, aturan gramatikal yang telah dirumuskan sebelumnya diterapkan untuk menghasilkan bentuk kata Bahasa Makassar yang sesuai secara morfologis. Aturan ini berfungsi untuk menghasilkan bentuk kata Bahasa Makassar yang sesuai secara morfologis.

Proses ini mencerminkan karakteristik utama RBMT, yaitu penggunaan aturan linguistik yang dirancang berdasarkan analisis bahasa. Dengan pendekatan ini, penerjemahan mampu menghasilkan terjemahan yang stabil meskipun sumber daya data relatif terbatas.

2.5.5 Pembentukan Kalimat

Apabila terdapat kata yang tidak ditemukan dalam kamus dan tidak dapat dianalisis secara morfologis, bentuk asli kata tersebut dipertahankan sebagai mekanisme cadangan untuk menjaga keterbacaan hasil terjemahan. Melalui seluruh tahapan ini, implementasi RBMT mampu menerjemahkan kalimat dari Bahasa Indonesia ke Bahasa Makassar secara otomatis dengan mengandalkan kamus Bahasa Indonesia–Bahasa Makassar serta aturan morfologi yang telah ditetapkan.

2.6 Pengujian

Pengujian dilakukan untuk mengevaluasi kinerja penerjemahan dalam mengubah teks dari Bahasa Indonesia ke Bahasa Makassar. Tujuan utama pengujian ini adalah untuk mengukur akurasi hasil terjemahan yang dihasilkan oleh implementasi RBMT. Pengujian ini menggunakan metode WER dan juga CER yang berfungsi untuk memberikan gambaran mengenai seberapa banyak kesalahan yang terjadi dalam proses penerjemahan dibandingkan dengan teks referensi yang benar.

Data uji yang digunakan dalam pengujian ini terdiri dari 50 kalimat dalam Bahasa Makassar yang diperoleh dari buku Tata Bahasa Makasar yang diterbitkan oleh Departemen Pendidikan dan Kebudayaan serta Percakapan Bahasa Makassar oleh Barbara Friberg dan Kembong Daeng. Kalimat-kalimat tersebut dipilih untuk mewakili berbagai aspek tata bahasa Makassar, mencakup berbagai konteks, mulai dari percakapan sehari-hari hingga kalimat dengan struktur yang lebih kompleks.

3. HASIL DAN PEMBAHASAN

3.1 Pengujian

Pengujian dilakukan untuk mengevaluasi kinerja penerjemahan dalam menghasilkan terjemahan yang sesuai dengan struktur dan kaidah Bahasa Makassar. Evaluasi dilakukan menggunakan dua metode, yaitu Word Error Rate (WER) dan Character Error Rate (CER). Kedua metode ini memberikan gambaran tingkat kesalahan pada hasil terjemahan, baik pada level kata maupun pada level karakter. Melalui pengujian ini, kualitas hasil penerjemahan dapat diukur secara objektif, sehingga dapat diketahui sejauh mana aturan dan kosakata yang digunakan mampu menghasilkan terjemahan yang akurat. Penilaian dilakukan menggunakan 50 kalimat uji yang telah dipilih berdasarkan frekuensi penggunaannya dalam percakapan sehari-hari. Hasil pengujian kemudian dianalisis untuk memperoleh nilai rata-rata yang mewakili performa keseluruhan penerjemahan.

3.1.1 WER

Pengujian Word Error Rate (WER) dilakukan untuk mengukur tingkat kesalahan pada hasil terjemahan berdasarkan perbandingan antara kata pada kalimat hasil terjemahan dengan kata pada kalimat acuan. Dalam penelitian ini, sebanyak 50 kalimat digunakan sebagai data uji untuk mengevaluasi performa penerjemahan pada berbagai jenis struktur kalimat, mulai dari kalimat sederhana hingga kalimat yang memiliki unsur morfologis tertentu. Metode WER dipilih karena mampu memberikan gambaran yang jelas mengenai tingkat kesalahan pada level kata, sehingga dapat menunjukkan sejauh mana hasil terjemahan yang benar secara keseluruhan. Semakin kecil nilai WER yang diperoleh, maka semakin sedikit kesalahan yang muncul, dan kualitas terjemahan dapat dikatakan semakin baik. Untuk memudahkan interpretasi nilai WER yang dihasilkan, penelitian ini menggunakan rentang penilaian sebagaimana ditunjukkan pada Tabel 3.

Tabel 3. Rentang Penilaian WER

Nomor	Score	Keterangan
1	0%	Terjemahan Sempurna
2	1% – 15 %	Sangat Baik
3	16% - 30%	Baik
4	31% – 45%	Cukup
5	46% - 60%	Kurang
6	> 60%	Buruk



Rentang penilaian tersebut digunakan sebagai acuan untuk mengklasifikasikan kualitas hasil terjemahan berdasarkan nilai WER yang diperoleh dari setiap kalimat uji. Pengujian dilakukan terhadap 50 kalimat uji yang disusun berdasarkan kalimat percakapan sehari-hari, sebagaimana ditunjukkan pada Tabel 4.

Tabel 4. Data Uji

Kalimat ke	Bahasa Indonesia	Bahasa Makassar
1	Apa itu?	Apa anjo?
2	Itu burung.	Anjo jangang-jangang.
3	Namanya Pak	
	Maming.	Arengna daeng Maming.
4	Baru tiga.	Nampai tallu.
...
50	Siapa temanmu tidur?	Inai nuagang tinro?

Setelah menghitung nilai WER pada semua 50 kalimat uji, diperoleh nilai WER seperti pada Tabel 5.

Tabel 5. Hasil Skor WER

Kalimat ke	Hasil Skor WER
1	0
2	0
3	0.33
....
50	0.33

Kalimat ke Hasil Skor WER

Untuk Menghitung hasil keseluruhan WER, digunakan nilai rata-rata sebagai berikut

$$Rata - Rata = \frac{\text{Jumlah Skor WER}}{\text{Jumlah Kalimat}} \equiv \frac{14.46}{50} = 0.289$$

Untuk memperoleh gambaran performa penerjemahan secara keseluruhan, nilai WER dihitung menggunakan nilai rata-rata dari seluruh kalimat uji. Berdasarkan perhitungan tersebut, diperoleh nilai WER keseluruhan sebesar 0,289 atau sekitar 28,9%.

Mengacu pada rentang penilaian WER yang telah ditetapkan, nilai WER sebesar 28,9% termasuk dalam kategori "Baik". Hasil ini menunjukkan bahwa penerjemahan berbasis aturan yang dikembangkan telah mampu menghasilkan terjemahan dengan tingkat kesalahan kata yang relatif rendah, meskipun masih terdapat kesalahan pada beberapa kalimat yang umumnya disebabkan oleh keterbatasan kosakata dalam kamus dan variasi struktur kalimat.

3.1.2 CER

Character Error Rate (CER) digunakan untuk mengukur tingkat kesalahan pada level karakter, sehingga memberikan gambaran seberapa dekat bentuk hasil terjemahan dengan bentuk kata referensi. Berbeda dengan WER yang berfokus pada kesalahan kata secara keseluruhan, CER memberikan toleransi terhadap kesalahan minor seperti perbedaan huruf atau ketidaktepatan morfologis yang tidak selalu mengubah makna kata secara signifikan. Pada penelitian ini, penghitungan CER dilakukan dengan membandingkan setiap karakter hasil terjemahan dengan karakter referensi pada seluruh kalimat uji. Sebagai acuan interpretasi nilai CER, digunakan rentang penilaian yang sama dengan WER sebagaimana ditunjukkan pada Tabel 6.

Tabel 6. Rentang Penilaian CER

Nomor	Score	Keterangan
1	0%	Terjemahan Sempurna
2	1% – 15 %	Sangat Baik
3	16% - 30%	Baik
4	31% – 45%	Cukup
5	46% - 60%	Kurang
6	> 60%	Buruk

Pengujian CER dilakukan terhadap 50 kalimat uji yang sama dengan pengujian WER, dan hasil perhitungan CER untuk setiap kalimat ditunjukkan pada Tabel 7.

Tabel 7. Hasil Skor CER

Kalimat ke	Hasil Skor CER
1	0



Kalimat ke	Hasil Skor CER
2	0
3	0.2
....
50	0.2

Untuk Menghitung hasil keseluruhan CER, digunakan nilai rata-rata sebagai berikut

$$Rata - Rata = \frac{\text{Jumlah Skor CER}}{\text{Jumlah Kalimat}} \equiv \frac{10.26}{50} = 0.21$$

Untuk memperoleh nilai CER keseluruhan, dilakukan perhitungan menggunakan nilai rata-rata dari seluruh kalimat uji. Berdasarkan perhitungan tersebut, diperoleh nilai CER akhir sebesar 0,21 atau 21%.

Berdasarkan rentang penilaian CER, nilai 21% termasuk dalam kategori "Baik". Nilai CER yang lebih rendah dibandingkan nilai WER menunjukkan bahwa secara umum bentuk karakter pada hasil terjemahan sudah mendekati bentuk referensi Bahasa Makassar. Hal ini mengindikasikan bahwa implementasi RBMT cukup efektif dalam menangani perubahan morfologis pada tingkat karakter, meskipun masih terdapat kesalahan yang berkaitan dengan pemilihan padanan kata atau struktur morfologi tertentu.

3.2 Analisis Hasil Penerjemahan

Hasil pengujian menunjukkan bahwa nilai Word Error Rate (WER) sebesar 0,289 dan Character Error Rate (CER) sebesar 0,21 berada pada kategori "Baik". Hasil penelitian ini sesuai dengan penelitian RBMT pada bahasa daerah lain, yang juga menunjukkan bahwa pendekatan berbasis aturan mampu menghasilkan tingkat kesalahan yang relatif rendah pada bahasa dengan keterbatasan sumber daya linguistik. Hal ini menunjukkan bahwa metode RBMT dapat digunakan pada bahasa daerah yang memiliki keterbatasan data.

3.2.1 Analisis Berdasarkan Word Error Rate (WER)

Word Error Rate (WER) digunakan untuk mengukur tingkat kesalahan penerjemahan pada level kata dengan membandingkan hasil terjemahan terhadap kalimat referensi. Nilai WER merepresentasikan proporsi kesalahan kata yang terjadi, sehingga semakin kecil nilai WER yang diperoleh, semakin sedikit kesalahan yang muncul pada hasil terjemahan.

Berdasarkan hasil pengujian, selain kalimat yang tidak mengandung kesalahan sama sekali atau yang nilainya 0 terdapat sejumlah kalimat dengan tingkat kesalahan rendah, yaitu dengan nilai WER berkisar antara 0,25 hingga 0,33. Kalimat-kalimat dengan nilai WER terendah pada kategori ini disajikan pada Tabel 8.

Tabel 8. Analisis Skor Terkecil WER

Nomor Kalimat	Refrensi	Hasi Terjemahan	S	D	I	N	Hasil
3	Arengna daeng Maming.	Arengna Angge Maming.	1	3	0.33		
6	Karaeng Abdul Muhammed.	Tatak Abdul Muhammed.	1	3	0.33		
11	Pammarianga ri Kajang.	Persawahan ri Kajang.	1	3	0.33		
38	Lima garang battalakna cincinna.	Lima garang battalak cincinna.	1	4	0.25		
39	Jai tau akboya jama-jamang.	Jai tau akboya paKaresoang.	1	4	0.25		

Kalimat-kalimat pada Tabel 8 menunjukkan bahwa kesalahan yang terjadi umumnya bersifat terbatas, yaitu berupa satu kesalahan kata pada setiap kalimat. Kesalahan tersebut biasanya muncul akibat perbedaan pemilihan kosakata, seperti penggunaan padanan kata yang tidak sepenuhnya identik dengan kalimat referensi, namun masih memiliki makna yang berdekatan.

Meskipun terdapat perbedaan pada satu kata, struktur kalimat secara keseluruhan tetap terjaga dan makna utama masih dapat dipahami dengan baik. Hal ini menunjukkan bahwa penerjemahan berbasis aturan mampu menghasilkan terjemahan yang cukup akurat, meskipun masih terdapat keterbatasan dalam pemilihan kosakata tertentu.

Sebagai perbandingan, kalimat-kalimat dengan nilai WER sebesar 0 yang tidak ditampilkan pada Tabel 9 umumnya merupakan kalimat dengan struktur sangat sederhana dan kosakata yang sepenuhnya tersedia dalam kamus dwibahasa, seperti "Apa anjo?", "Anjo jangang-jangang.", dan "Inai arengna?". Pada kondisi tersebut, hasil terjemahan identik dengan kalimat referensi karena tidak diperlukan transformasi leksikal yang kompleks.

Selain nilai WER terendah, hasil pengujian juga menunjukkan adanya kalimat dengan nilai WER yang lebih tinggi. Nilai WER 0,67 banyak ditemukan pada kalimat yang mengalami perbedaan pemilihan kata, khususnya pada kata benda atau kata kerja tertentu. Kesalahan pada kategori ini umumnya disebabkan oleh keterbatasan kosakata dalam kamus serta adanya variasi leksikal yang belum sepenuhnya tertangani oleh aturan yang tersedia. Kalimat-kalimat dengan nilai WER tertinggi ditunjukkan pada Tabel 9.

Tabel 9. Analisis Skor Tertinggi WER

Nomor Kalimat	Refrensi	Hasi Terjemahan	S	D	I	N	Hasil
18	Ri allonna Kammisika.	Ri allo Kammisik.	2	3	0.67		
32	Napangalleangi anakna taipa.	Dia mangalleang anakna taipa.	1	1	0.67		



Nomor Kalimat	Refrensi	Hasi Terjemahan	S	D	I	N	Hasil
33	Jai tau mange ri ballak lomppa.	Jai tau mange ke istana.	3	1	6	0.67	

Berdasarkan Tabel 9, nilai WER tertinggi yang diperoleh adalah sebesar 0,67. Kalimat-kalimat dengan nilai tersebut umumnya memiliki struktur yang lebih kompleks dan panjang, serta mengandung kombinasi beberapa unsur morfologis dan variasi struktur sintaksis. Pada kondisi ini, penerjemahan berbasis kata per kata belum sepenuhnya mampu merepresentasikan hubungan antarkata secara utuh, sehingga menghasilkan perbedaan yang lebih besar dibandingkan dengan kalimat referensi.

3.2.2 Analisis Berdasarkan Character Error Rate (CER)

Analisis Character Error Rate (CER) dilakukan untuk mengukur tingkat kesalahan penerjemahan pada level karakter. Metode ini digunakan untuk mengevaluasi sejauh mana kesesuaian bentuk kata hasil terjemahan dibandingkan dengan kalimat referensi, sehingga dapat memberikan gambaran ketepatan penerapan aturan morfologi dan pembentukan karakter dalam sistem penerjemahan. Berdasarkan hasil pengujian, diperoleh nilai rata-rata CER sebesar 0,21, yang menunjukkan bahwa secara umum hasil terjemahan memiliki tingkat kemiripan karakter yang cukup tinggi dengan kalimat referensi.

Tabel 10 menyajikan beberapa contoh kalimat dengan nilai CER terkecil. Pada Tabel 10 terlihat bahwa sebagian besar kalimat hanya mengalami sedikit atau bahkan tidak mengalami kesalahan karakter sama sekali.

Tabel 10. Analisis Skor Terkecil CER

Nomor Kalimat	Refrensi	Hasi Terjemahan	S	D	I	N	Hasil
10	Garring apai?	Garring apa?	0	1	13	0.08	
17	Tallu meterek kaeng naballu ri tokoa.	Tallu meterek kaeng riballi ri toko.	1	1	37	0.05	
31	Taipa nasambila andikku.	Taipa risambila andikku.	2		24	0.08	
38	Lima garang battalakna cincinna.	Lima garang battalak cincinna.	0	2	32	0.06	
43	Wah, gakgana bajunna.	Wah, gakgana bajuna.	0	1	21	0.05	

Sebagaimana ditunjukkan pada Tabel 10, nilai CER yang rendah atau mendekati 0 menandakan bahwa hasil terjemahan memiliki tingkat kesesuaian karakter yang sangat tinggi dengan kalimat referensi. Kondisi ini umumnya ditemukan pada kalimat-kalimat sederhana serta kalimat yang mengikuti pola morfologi yang konsisten. Kesalahan yang muncul pada kelompok ini sebagian besar berupa perubahan satu karakter, penghilangan karakter tertentu, atau variasi ejaan minor yang tidak berdampak signifikan terhadap makna kata. Hal tersebut menunjukkan bahwa sistem penerjemahan mampu mempertahankan struktur karakter dengan baik, serta aturan perubahan awalan dan akhiran yang diterapkan telah bekerja secara efektif pada sebagian besar data uji.

Sebaliknya, Tabel 11 menampilkan contoh kalimat dengan nilai CER terbesar, yang menunjukkan adanya tingkat perbedaan karakter yang lebih tinggi antara hasil terjemahan dan kalimat referensi. Peningkatan nilai CER pada kelompok ini umumnya disebabkan oleh perbedaan pemilihan kosakata, penggunaan padanan kata yang berbeda secara leksikal, atau adanya reformulasi kata dalam proses penerjemahan. Akibatnya, jumlah operasi substitusi, penghapusan, dan penyisipan karakter menjadi lebih besar, sehingga berdampak langsung pada nilai CER yang dihasilkan.

Tabel 11. Analisis Skor Terbesar CER

Nomor Kalimat	Refrensi	Hasi Terjemahan	S	D	I	N	Hasil
4	Nampai tallu.	Beru tallu.	4	2	13	0.46	
9	Sallo kamma.	Sallo sikali.	5		1	12	0.50
29	Jai rupanna; akkulie kondo, jangang, kitik, jukuk.	Bermacam-macam; biasa kondo, jangang, kitik, ikan.	12	11	10	51	0.65
32	Napangalleangi anakna taipa.	Dia angalleang anakna taipa.	3	11	10	28	0.86
47	Arragadak, basai baju beruna.	Amboi, jakmerek baju beruna.	8	4	3	29	0.52
32	Napangalleangi anakna taipa.	Dia angalleang anakna taipa.	3	11	10	28	0.86

Berdasarkan Tabel 11, nilai CER menengah hingga tinggi umumnya muncul pada kalimat yang mengalami perbedaan kosakata yang cukup signifikan atau adanya substitusi kata dengan padanan yang berbeda secara leksikal. Pada kasus ini, meskipun struktur kalimat masih dapat dipahami, perbedaan pilihan kata menyebabkan peningkatan jumlah operasi substitusi, penghapusan, dan penyisipan karakter. Namun demikian, nilai CER pada kalimat-kalimat tersebut masih cenderung lebih rendah dibandingkan nilai Word Error Rate (WER), yang mengindikasikan bahwa kesalahan lebih dominan terjadi pada pemilihan kata, bukan pada pembentukan karakter atau struktur morfologis secara keseluruhan.

4. KESIMPULAN

Penelitian ini membuktikan bahwa metode *Rule Based Machine Translation* (RBMT) dapat diterapkan secara efektif untuk menerjemahkan teks dari Bahasa Indonesia ke Bahasa Makassar dengan tingkat akurasi yang tergolong baik. Penerjemahan diimplementasikan dengan memanfaatkan kamus dwibahasa serta aturan morfologi yang mencakup



transformasi awalan dan akhiran, sehingga mampu menghasilkan terjemahan yang sesuai dengan struktur dasar dan karakteristik linguistik Bahasa Makassar. Aturan perubahan bentuk kata yang dirumuskan dalam penelitian ini disusun secara teratur dan dapat diterapkan secara konsisten, serta tidak hanya mengubah awalan dan akhiran, tetapi juga menyesuaikan bentuk kata dasar, seperti penggandaan huruf awal tertentu, agar sesuai dengan aturan morfologi Bahasa Makassar. Hasil evaluasi pengujian terjemahan menunjukkan nilai rata-rata *Word Error Rate* (WER) sebesar 0,289 dan evaluasi pengujian terjemahan *Character Error Rate* (CER) sebesar 0,21, yang mengindikasikan bahwa tingkat kesalahan terjemahan relatif rendah baik pada kata maupun pada karakter. Secara fungsional, penerjemahan bekerja optimal pada kosakata yang telah terakomodasi dalam kamus dan pola morfologi yang telah dirumuskan, namun masih menghadapi keterbatasan pada penanganan kata bermakna ganda, kosakata yang belum terdaftar, serta kalimat dengan struktur sintaksis yang lebih kompleks. Selain itu, keterbatasan implementasi terjemahan dalam menangani variasi konteks dan gaya bahasa menunjukkan perlunya pengembangan aturan sintaksis yang lebih mendalam. Oleh karena itu, penelitian selanjutnya disarankan untuk memperluas cakupan kosa kata, memperkaya dan memformalkan aturan morfologi, serta mengintegrasikan mekanisme ambiguitas makna guna meningkatkan kualitas dan cakupan hasil terjemahan, sekaligus mendukung upaya pelestarian Bahasa Makassar melalui pendekatan komputasional yang berkelanjutan.

REFERENCES

- [1] S. Dinanta and B. Ginting, "Strategi Penerjemahan Dalam Bahasa Indonesia Politeknik Negeri Lhokseumawe," *Jurnal Bahasa*, vol. 11, no. 1, p. 163, Mar. 2022, doi: 10.24114/kjb.v11i1.33504.
- [2] N. Sabrina, "Pergeseran Makna Dalam Penerjemahan dari Bahasa Inggris Ke Dalam Bahasa Indonesia Pada teks Perjanjian Internasional Di Bidang Pertahanan," *Jurnal Linguistik Terapan*, vol. 5, no. 2, pp. 30–38, Nov. 2015.
- [3] Y. Galingga and G. Tambunsaribu, "Penerjemahan Idiomatis Peter Newmark dan Mildred Larson," *Dialektika : Jurnal Bahasa, Sastra, dan Budaya*, vol. 8, no. 1, pp. 57–70, Jun. 2021, doi: 10.33541/dia.v8i1.3112.
- [4] A. W. Mahastama, "Model Berbasis Aturan Untuk Transliterasi Bahasa Jawa Dengan Aksara Latin Ke Aksara Jawa," *Jurnal Buana Informatika*, vol. 13, no. 2, pp. 146–154, Oct. 2022, doi: 10.24002/jbi.v13i02.6526.
- [5] N. Khasanah and A. Subiyanto, "Syllable Structure and Nasalization Change in Makassar Language," *Lingua Cultura*, vol. 14, no. 1, pp. 107–113, Jul. 2020, doi: 10.21512/lc.v14i1.6357.
- [6] M. H. Dr. Asriani Abbas, *Morfosintaksis Bahasa Makassar*, vol. 1. Malang: CV. Rena Cipta Mandiri, 2021.
- [7] L. Putri, "Analisis Kontrastif Bahasa Makassar dan Bahasa Indonesia dalam Proses Afiksasi," *Metafora: Jurnal Pembelajaran Bahasa Dan Sastra*, vol. 11, no. 1, pp. 183–190, Apr. 2024, doi: 10.30595/mtf.v11i1.21135.
- [8] H. Wang, H. Wu, Z. He, L. Huang, and K. W. Church, "Progress in Machine Translation," *Engineering*, vol. 18, pp. 143–153, Nov. 2022, doi: 10.1016/j.eng.2021.03.023.
- [9] M. S. Alam and A. A. Suryani, "Minang and Indonesian Phrase-Based Statistical Machine Translation," *Journal of Informatics and Telecommunication Engineering*, vol. 5, no. 1, pp. 216–224, Jul. 2021, doi: 10.31289/jite.v5i1.5308.
- [10] F. Abubekar Tau, P. Aliandu, and E. Ngaga, "Aplikasi Penerjemahan Bahasa Indonesia-Bahasa Ende Lio Menggunakan Metode Rule Based Berbasis Web," *Jurnal Ilmu Komputer Universitas Katolik Widya Mandira*, vol. 1, no. 1, pp. 11–16, May 2023.
- [11] R. Alex Richo, A. Asroni, G. Indrawan, and L. Joni Erawati Dewi, "Pengembangan Transliterasi Offline Teks Aksara Bali dengan Pendekatan Metode Rule-Based Berdasarkan Unicode," *Resistor*, vol. 6, no. 3, pp. 138–144, Dec. 2023, doi: 10.31598/jurnalresistor.v6i3.1485.
- [12] J. Coleman, B. Krishnamachari, K. Iskarous, and R. Rosales, "LLM-Assisted Rule Based Machine Translation for Low/No-Resource Languages," May 2024, pp. 67–87. doi: 10.18653/v1/2024.americasnlp-1.9.
- [13] D. Torregrosa *et al.*, "Aspects of Terminological and Named Entity Knowledge within Rule-Based Machine Translation Models for Under-Resourced Neural Machine Translation Scenarios," Sep. 2020, doi: <https://doi.org/10.48550/arXiv.2009.13398>.
- [14] J. Nitzke and S. Hansen-Schirra, *A short guide to postediting Translation and Multilingual Natural Language Processing* 16, vol. 16. Berlin: Language Science Press, 2021. doi: 10.5281/zenodo.5646896.
- [15] F. Fadhlia, "Designing a Website Based Aceh Translation Application Using Rule-Based," *International Journal Education and Computer Studies*, vol. 1, no. 1, pp. 16–23, May 2021, doi: 10.35870/ijecs.v1i1.330.
- [16] M. A. Sghaier and M. Zrigui, "Rule-based machine translation from Tunisian dialect to modern standard Arabic," *Procedia Comput Sci*, vol. 176, pp. 310–319, 2020, doi: 10.1016/j.procs.2020.08.033.
- [17] M. Aditya, "Rancang Bangun Aplikasi Web untuk Penerjemahan Otomatis Bahasa Indonesia dan Bahasa Minangkabau dengan Fitur Text to Speech," Pekanbaru, Jun. 2025.
- [18] R. Hermansyah, "Implementasi Metode Rule Based Machine Translation dalam Proses Penerjemahan Teks Bahasa Indonesia ke Bahasa Alas Tenggara Tugas Akhir," Pekanbaru, Sep. 2024.
- [19] K. Indriani, L. G. Da Silva, and S. Sumanto, "Implementasi Penerjemah Bahasa Tetun Indonesia (Timor Leste Tetun) Berbasis Web," *Applied Information System and Management (AISM)*, vol. 4, no. 1, pp. 51–55, Apr. 2021, doi: 10.15408/aism.v4i1.20273.
- [20] S. Lee *et al.*, "A Survey on Evaluation Metrics for Machine Translation," *Mathematics*, vol. 11, no. 4, pp. 2–22, Feb. 2023, doi: 10.3390/math11041006.
- [21] M. N. Darpito, K. Firdausy, and A. Fadlil, "Perbandingan Unjuk Kerja Library Optical Character Recognition (OCR) dalam Pengenalan Teks pada Dokumen Digital," *Jurnal Informatika Polinema*, vol. 11, no. 3, pp. 273–282, May 2025, doi: 10.33795/jip.v11i3.7025.