

BAB III

METODOLOGI PENELITIAN

Bab ini membahas tentang metode penelitian yang digunakan untuk menjawab pertanyaan penelitian. Pembahasan yang dicakup adalah desain penelitian secara menyeluruh, sumber data penelitian, metode pengumpulan data dan instrumen penelitian, dan metode analisis data.

3.1 Desain Penelitian

Fokus dari penelitian ini menilai kualitas hasil terjemahan *Google Translate* pada teks akademik dari Bahasa Indonesia ke dalam Bahasa Inggris. Kualitas tersebut dinilai dari tiga sudut pandang. Sudut pandang pertama berasal dari peneliti sendiri yang mengkaji terjemahan dari segi struktur dan makna. Sudut pandang kedua berasal dari pengguna *GT*. Mereka menilai kualitas terjemahan *GT* berkenaan dengan aspek keberterimaan dan keterbacaan. Kemudian sudut pandang ketiga berasal dari *rater* yang menilai kualitas *GT* pada tiga aspek, yaitu keakuratan, keberterimaan, dan keterbacaan.

Untuk mencapai tujuan penelitian tersebut, metode yang dipakai adalah dengan menerapkan pendekatan campuran, atau yang disebut dengan *mixed method*. Menurut Creswell (2014) *mixed method* secara garis besar adalah pendekatan penelitian dengan mengkombinasikan metode kualitatif dengan kuantitatif. Pendekatan kualitatif dilakukan karena data berupa teks dan analisis terhadap kualitas terjemahan dari sudut pandang peneliti dilakukan secara kualitatif dan eksploratif dengan membandingkan teks sumber dan teks terjemahan. Perbandingan dilakukan dengan melakukan kajian dari segi struktur dan makna dalam kedua versi teks. Pendekatan kuantitatif dilakukan karena data yang diperoleh dari pengguna dan *rater* adalah berupa skor penilaian terjemahan dan dianalisis secara kuantitatif dan kualitatif.

3.2 Sumber Data

Data untuk penelitian terdiri dari teks dalam Bahasa Indonesia, teks terjemahan dalam Bahasa Inggris, penilaian terjemahan dari pengguna *GT* dan penilaian *GT* dari *rater*.

Menik Winiharti, 2022

KUALITAS TERJEMAHAN GOOGLE TRANSLATE PADA TEKS AKADEMIK: KAJIAN STRUKTUR DAN MAKNA

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

3.2.1 Teks

Data utama untuk penelitian ini berupa teks akademik yang ditulis oleh akademisi Indonesia dengan afiliasi perguruan tinggi di Indonesia. Teks tersebut merupakan teks publikasi ilmiah yang telah diterbitkan oleh jurnal nasional di Indonesia. Teks artikel ini dicari dan diunduh melalui laman <https://www.neliti.com/id/journals>. Pilihan jatuh pada laman ini dikarenakan banyak sekali artikel ilmiah yang dapat diunduh secara gratis di sini, dan sudah diterbitkan di jurnal nasional maupun internasional. Tiga bidang ilmu pengguna yang dipilih adalah bidang ilmu ekonomi, informatika, dan bahasa/sastra, yang masing-masing dapat mewakili ilmu sosial, sains, dan humaniora berturut-turut. Oleh karena itu teks yang dikumpulkan pun berasal dari tiga bidang ilmu ini, dengan menyaring pilihan bidang studi di laman tersebut. Lima artikel dikumpulkan dari tiap bidang ilmu, sehingga seluruhnya terkumpul limabelas artikel.

Penulis kelimabelas artikel tersebut tidak diketahui apakah menggunakan GT atau tidak karena data yang diambil untuk keperluan studi ini hanyalah teks asli dalam Bahasa Indonesia yang mereka tulis.

3.2.2 Pendapat Pengguna

Pengguna memiliki peran penting dalam penerjemahan karena merekalah yang menggunakan dan menikmati produk akhir dari proses penerjemahan. Karena itu penelitian ini mengumpulkan data dari pengguna terkait pendapat mereka terhadap aspek keberterimaan dan keterbacaan terjemahan GT. Pengguna ini terdiri dari mahasiswa dan dosen S1 yang berasal dari tiga bidang ilmu, yaitu ekonomi, informatika dan bahasa/sastra. Jumlah pengguna yang menjadi responden survei tersaji dalam tabel 3.1 berikut.

Tabel 3.1 Data Pengguna

Bidang Ilmu	Mahasiswa S1 (Semester 2-6)	Dosen S1	Jumlah
Ekonomi	30	5	35
Informatika	30	5	35
Bahasa/Sastra	30	5	35
Jumlah	90	15	105

Pada awalnya pengguna ini dipilih secara acak dengan mempertimbangkan faktor kemudahan untuk menghubungi mereka melalui dosen masing-masing. Teknik pengambilan sampel semacam ini disebut dengan *convenience sampling* yang merupakan “a type of nonprobability sampling in which people are sampled simply because they are ‘convenient’ sources of data for researchers.” (Battaglia, 2008).

Meskipun demikian untuk mendapatkan penilaian yang tepat dari pengguna maka skor TOEFL terakhir mereka digunakan sebagai ukuran kemampuan mereka karena hanya yang memiliki skor minimal 501 yang diperhitungkan pendapatnya. Menurut Dewi dkk. (2015) dan Raharjo (2020) skor TOEFL dapat dianggap sebagai salah satu alat ukur akan kompetensi berbahasa-Inggris seseorang. Berikut adalah gambaran kemampuan pengguna dengan skor TOEFL IBT dikonversikan ke dalam versi PBT:

- 500 – 586: *competent*
- 587 – 636: *good*
- 637 – 666: *very good*
- 667 – 677: *expert*

(disarikan dari:

https://www.ets.org/s/toefl/pdf/linking_toefl_ibt_scores_to_ielts_scores.pdf)

Universitas Pendidikan Indonesia Bandung pun mensyaratkan mahasiswanya untuk “memiliki kemampuan dalam salah satu bahasa resmi PBB, yang dibuktikan dengan hasil tes terstandar”. Salah satu tes yang diakui sebagai alat ukur kompetensi Bahasa Inggris oleh UPI adalah TOEFL. (*Pedoman Penyelenggaraan Pendidikan Universitas Pendidikan Indonesia*, 2020, hlm. 100-101) Untuk penelitian ini TOEFL yang dimaksud adalah yang menyerupai TOEFL atau disebut dengan *TOEFL equivalent* atau *TOEFL-like*.

Setelah skor TOEFL mereka diketahui, maka dari masing-masing bidang ilmu seluruh pengguna disaring hingga mendapatkan 35 pengguna sehingga seluruhnya melibatkan 105 pengguna.

Sebagai informasi tambahan, pengguna yang dimaksud dalam studi ini tidak berhubungan sama sekali dengan penulis artikel yang teksnya diambil sebagai data BS. Pengguna hanya menilai hasil terjemahan GT yang data BSnya diambil dari teks yang ditulis oleh penulis artikel tersebut.

Menik Winiharti, 2022

KUALITAS TERJEMAHAN GOOGLE TRANSLATE PADA TEKS AKADEMIK: KAJIAN STRUKTUR DAN MAKNA

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

3.2.3 Pendapat *Rater*

Larson (1998) menyatakan bahwa akan bagus bila sebuah produk terjemahan dites, dalam arti dinilai kualitasnya. Pengetes harus memahami prinsip-prinsip penerjemahan dan berpengetahuan baik terutama dalam bahasa target. Akan lebih bagus jika ada reviewer yang berpengetahuan dalam konsep penerjemahan idiomatik, dan memiliki rasa yang bagus dalam dalam hal *style* dan struktur tata bahasa.

Salah satu tujuan penelitian ini adalah mengidentifikasi pendapat *rater* terkait aspek keakuratan, keberterimaan, dan keterbacaan terjemahan. Untuk dapat menilai ketiga aspek tersebut, rater ini dipilih dari mereka yang memiliki kualifikasi tertentu, yaitu:

- a. Menguasai Bahasa Indonesia dan Bahasa Inggris dengan sangat baik.
- b. Berlatarbelakang pendidikan S2 dalam ilmu Bahasa Inggris, dan/atau berkecimpung dalam dunia penerjemahan dalam waktu cukup lama.
- c. Dosen S1/S2 jurusan Bahasa Inggris
- d. Berkecimpung dalam dunia penerjemahan, baik sebagai penerjemah biasa maupun profesional.
- e. Reviewer jurnal nasional atau internasional menjadi nilai tambah.

Sejumlah tujuh *rater* terlibat dalam survei penelitian ini. Dari kriteria yang telah ditetapkan di atas, keenam *rater* memenuhi semua kriteria a-d, dan hanya satu *rater* berlatar belakang pendidikan S1 namun memiliki pengalaman cukup lama dalam bidang penerjemahan, baik sebagai profesional maupun memberikan kelas-kelas penerjemahan.

Proses pemilihan para *rater* dilakukan dengan pendekatan personal karena saya mengenal ketujuh *rater* ini sebagai kolega. Secara personal pula saya menjelaskan kepada mereka tentang survei yang harus mereka isi dan bagaimana konsep keakuratan, keberterimaan, dan keterbacaan dari sebuah terjemahan.

3.3 Metode Pengumpulan Data

Langkah pertama dalam mengumpulkan data adalah dengan mendapatkan artikel publikasi ilmiah dari laman <https://www.neliti.com/id/journals>. Spesifikasi artikel

Menik Winiharti, 2022

KUALITAS TERJEMAHAN GOOGLE TRANSLATE PADA TEKS AKADEMIK: KAJIAN STRUKTUR DAN MAKNA

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

yang diunduh dari laman ini dibahas di bagian 3.2.1. Sebanyak limabelas artikel dari tiga bidang studi dikumpulkan kemudian disimpan secara elektronik dalam bentuk teks.

Langkah berikutnya adalah membaca teks sumber yang tertulis dalam Bahasa Indonesia, diikuti dengan menseleksi data. Kalimat sebagai data dalam bahasa sumber diputuskan dipilih dari bagian ‘kesimpulan’ dengan alasan bahwa pada bagian kesimpulan umumnya penulis atau peneliti menuliskan pemikiran mereka berdasarkan hasil pembahasan dalam bagian analisis, sehingga mereka cenderung menggunakan kata-kata mereka sendiri untuk merangkum temuan penelitian secara menyeluruh. Kemudian seluruh kalimat dalam bagian kesimpulan di-copy dan *paste* ke dalam Microsoft Word. Setelah itu dilakukan penyaringan untuk menetapkan kalimat yang akan diterjemahkan melalui GT. Akhirnya dari satu bagian kesimpulan diseleksi menjadi empat kalimat pertama, sehingga dari limabelas artikel terkumpul enampuluh kalimat dengan masing-masing bidang ilmu terkumpul duapuluh kalimat. Jumlah ini cukup untuk menjadi bahan analisis dengan pemahaman bahwa ukuran sampel yang layak dalam sebuah penelitian paling tidak berjumlah tigapuluh (Sugiyono, 2019).

3.3.1 Pengumpulan Data Teks Terjemahan

Kalimat Bahasa Indonesia yang telah disimpan dalam file MS Word diperiksa kembali untuk memastikan bahwa data tersebut sudah sesuai dengan bidang ilmu yang dipilih sehingga dikatakan bisa mewakili tiap bidang ilmu. Setelah itu tiap kalimat di-copy dan *paste* ke dalam laman *GT* <https://translate.google.com/> untuk diperoleh terjemahannya dalam Bahasa Inggris. Tahapan ini dilaksanakan pada bulan Februari 2020. Kemudian terjemahan yang dihasilkan *GT* ini dipindah ke kolom di samping kalimat sumbernya untuk disimpan dalam satu file. Dengan cara ini kalimat sumber dan kalimat terjemahannya terlihat berdampingan secara jelas dalam satu tabel sehingga memudahkan analisis.

3.3.2 Pengumpulan Data dari Pengguna

Untuk mendapatkan data yang berupa penilaian pengguna, survei untuk pengguna disusun kemudian disebarakan kepada mereka secara daring. Dalam survei ini pengguna diminta untuk menilai duapuluh kalimat dan terjemahannya dalam

Menik Winiharti, 2022

KUALITAS TERJEMAHAN GOOGLE TRANSLATE PADA TEKS AKADEMIK: KAJIAN STRUKTUR DAN MAKNA

Universitas Pendidikan Indonesia | repository.upi.edu | perpustakaan.upi.edu

Bahasa Inggris dari aspek keberterimaan dan keterbacaan terjemahan. Survei didistribusikan secara daring pada periode Maret – Desember 2020. Disain survei beserta rubriknya dibahas di subbab 3.4 di bawah.

3.3.3 Pengumpulan Data dari *Rater*

Untuk penilaian dari *rater*, survei yang berbeda disusun karena mengandung tiga aspek penilaian terjemahan, yaitu keakuratan, keberterimaan, dan keterbacaan. Penilaian dari *rater* ini sedikit berbeda dengan penilaian dari pengguna karena ada satu aspek yang ditambahkan kepada *rater*, yaitu keakuratan. *Rater* dianggap memiliki kualifikasi menilai tingkat keakuratan sebuah terjemahan dengan latar belakang pendidikan yang mereka miliki dan profesi yang mereka geluti. Survei terhadap *rater* didistribusikan secara daring pada periode April – Desember 2020. Disain survei dan rubrik yang diterapkan untuk *rater* disajikan dalam subbab 3.4 di bawah.

3.3.4 Wawancara kepada *Rater*

Sugiyono (2019, 229-33) menyatakan bahwa wawancara adalah salah satu teknik pengumpulan data dengan tujuan memperoleh informasi dari responden hal-hal yang lebih mendalam dengan jumlah responden yang sedikit.

Mengingat bahwa *rater* merupakan orang yang ahli di bidang bahasa dan/atau penerjemahan, maka mereka diminta untuk menjawab secara kualitatif pertanyaan lanjutan yang diajukan kepada mereka. Jawaban mereka diperlukan untuk mendapatkan informasi tambahan dari mereka terkait kualitas terjemahan GT, sekaligus memperkaya hasil penelitian ini secara keseluruhan. Wawancara ini dilakukan secara tertulis setelah *rater* mengisi survei seperti yang dibahas di subbagian sebelumnya. Pertanyaan berjenis *open-ended* dengan tujuan mendapatkan jawaban yang luas dan tidak membatasi pendapat *rater*. Connor Desai & Reimers, (2019) menyatakan bahwa responden memberikan data berkualitas tinggi untuk pertanyaan terbuka. Mereka menemukan bahwa responden dengan tekun menjawab pertanyaan serta menjawab lebih dari yang diminta untuk memberikan jawaban lengkap.

Ada tiga pertanyaan yang diajukan kepada mereka, yaitu:

1. Menurut Anda, apa saja kriteria atau indikator sebuah terjemahan yang tepat?
2. Bagaimana sebaiknya langkah-langkah menerjemahkan yang tepat?
3. Menurut Anda, bagaimana hasil terjemahan *google* di atas (dalam survei) secara keseluruhan?

3.4 Instrumen Penelitian

Survei disusun untuk mendapatkan data dari pengguna dan rater. Karena pengguna berasal dari tiga bidang ilmu yang berbeda, maka disusun tiga macam survei untuk pengguna dari tiga kelompok tersebut. Perbedaan dalam tiga survei itu adalah pada kalimat-kalimat yang ditanyakan. Pengguna dari bidang ilmu ekonomi misalnya harus familiar dengan kalimat yang ditanyakan, sehingga kalimat yang ditanyakan kepada mereka pun harus diambil dari artikel ekonomi. Dengan demikian pengguna dari tiap bidang ilmu diasumsikan menghadapi kalimat yang sesuai bidang mereka masing-masing. Persamaan dari ketiga survei tersebut adalah pada aspek penilaiannya, di mana kesemuanya diminta untuk menilai aspek keberterimaan dan keterbacaan terjemahan. Dalam tiap survei ada duapuluh kalimat yang ditanyakan, sehingga keenampuluh kalimat data tersebar semua kepada pengguna.

Kemudian survei untuk *rater* disusun tersendiri karena ada satu aspek keakuratan yang dinilai, di mana aspek ini tidak ikut dinilai oleh pengguna. Selain itu duapuluh kalimat yang ditanyakan kepada *rater* merupakan gabungan dari kalimat-kalimat yang berasal dari tiga bidang ilmu, sehingga penelitian ini merancang empat survei untuk empat macam responden.

3.4.1 Rubrik Penilaian dalam Survei

Di dalam masing-masing survei terdapat rubrik penilaian terjemahan untuk memudahkan responden menilai terjemahan dan dengan tujuan mereka memiliki satu acuan yang sama dalam melakukan penilaian. Rubrik penilaian ini khusus dirancang untuk penelitian ini berdasarkan rubrik penilaian yang diajukan oleh Carroll (1966); Lai (2011); Liu (2005); dan Nababan dkk., (2012). (Lihat tabel 2.1 sampai tabel 2.8 di hlm. 51-58).

Aspek penilaian terjemahan yang dipakai untuk penelitian ini ada tiga yaitu keakuratan, keberterimaan, dan keterbacaan. Rubrik yang dikembangkan untuk masing-masing aspek tersebut disajikan dan dibahas secara rinci di bab II hlm. 59 – 61 di mana tabel 2.9 menyajikan aspek keakuratan, tabel 2.10 menyajikan aspek keberterimaan, dan tabel 2.11 menyajikan aspek keterbacaan.

Rubrik-rubrik tersebut menjadi petunjuk *rater* untuk menilai kualitas terjemahan GT dipandang dari aspek keakuratan, keberterimaan, dan keterbacaan berturut-turut. Sementara itu rubrik yang diterapkan kepada pengguna hanyalah rubrik keberterimaan dan rubrik keterbacaan.

Berikut contoh bentuk pertanyaan yang diajukan kepada pengguna.

Kalimat 7

Aplikasi sistem pakar “Awas Meningitis!” dapat melakukan proses diagnosis penyakit sesuai data rekomendasi yang didapatkan dari pakar.

Application of the expert system "Beware Meningitis!" can carry out the disease diagnosis process according to the recommendation data obtained from experts.

Penilaian:

Kalimat 7	1	2	3	4	5
Keberterimaan					
Keterbacaan					

Saran:

Kemudian contoh pertanyaan yang diajukan kepada rater adalah sebagai berikut.

Kalimat 17

Hal ini menunjukkan bahwa teks terjemahan itu bersifat romantis dengan mengunggulkan manusia, cita-cita pribadi, pikiran, dan dunia idealnya sehingga meniadakan dunia sosialnya.

This shows that the translated text is romantic by favoring humans, personal ideals, thoughts, and the ideal world so that it negates the social world.

Penilaian:

Kalimat 17	1	2	3	4	5
Keakuratan					
Keberterimaan					
Keterbacaan					

Saran:

3.5 Metode Analisis Data

Setelah data terjemahan terkumpul dari GT, maka analisis yang pertama dilakukan adalah dengan mengelompokkan data menurut jenis kalimat bahasa sumber. Pengelompokkan ini dilakukan berdasarkan teori tentang jenis kalimat menurut jumlah klausanya yang disampaikan oleh Moeliono dkk. (2017). Setelah dikelompokkan, tiap kalimat dianalisis dengan melakukan perbandingan antara teks sumber dan teks target, berdasarkan struktur dan maknanya. Di akhir analisis tiap kalimat ditentukan tingkat keakuratan dan keberterimaan dengan mengacu pada rubrik yang sudah ditetapkan. Tingkat keakuratan diperoleh setelah melakukan analisis makna, sementara tingkat keberterimaan setelah melakukan analisis pada struktur kalimat terjemahan. Dalam melakukan analisis struktur dan makna terhadap dua versi teks, identifikasi masalah yang muncul dalam terjemahan terkait struktur dan makna pun dilakukan untuk kemudian dapat mengelompokkan permasalahan itu berdasarkan jenis-jenisnya. Di samping itu, dari hasil melakukan analisis di tahap ini, identifikasi pada strategi penerjemahan juga dilakukan untuk menetapkan tipe strategi yang diterapkan GT.

Langkah berikutnya adalah menghitung pendapat pengguna terhadap tiap kalimat terkait keberterimaan dan keterbacaan terjemahan. Karena pendapat pengguna dinyatakan dalam angka atau skor, maka rata-rata penilaian dari seluruh pengguna terhadap keberterimaan dan keterbacaan sebuah kalimat bisa dilakukan.

Rata-rata skor penilaian dihitung dengan cara yang umum, yaitu dengan menjumlah skor sebuah kalimat dari 35 pengguna lalu dibagi dengan 35.

$$\text{Rata-rata keberterimaan satu kalimat} = \frac{\text{Jumlah 35 skor keberterimaan}}{35}$$

$$\text{Rata-rata keterbacaan satu kalimat} = \frac{\text{Jumlah 35 skor keterbacaan}}{35}$$

Setelah mendapatkan skor rata-rata terhadap keberterimaan dan keterbacaan dari pengguna, yang dilakukan selanjutnya adalah menganalisis angka-angka itu dengan mengelompokkannya ke dalam interval skor (disajikan pada sub-bab 3.6) untuk menetapkan sebuah terjemahan termasuk dalam kelompok yang mana. Kemudian terjemahan dianalisis secara kualitatif dari segi struktur dan maknanya untuk kemudian dibandingkan dengan ulasan dan nilai yang diberikan peneliti. Contoh kalimat yang dibahas di sini diambil dari terjemahan yang mendapatkan skor tertinggi dan terendah.

Setelah penghitungan dan analisis pada pendapat pengguna selesai dilakukan, langkah berikutnya adalah melakukan analisis pada pendapat kelompok rater. Sama halnya dengan yang dilakukan pada pendapat pengguna, rata-rata skor yang diberikan rater juga dihitung, namun karena melibatkan tiga aspek penilaian maka penghitungan pun dilakukan pada tiga aspek itu: keakuratan, keberterimaan, dan keterbacaan. Karena ada tujuh rater maka pembagiannya pun 7.

$$\text{Rata-rata keakuratan satu kalimat} = \frac{\text{Jumlah 7 skor keakuratan}}{7}$$

$$\text{Rata-rata keberterimaan satu kalimat} = \frac{\text{Jumlah 7 skor keberterimaan}}{7}$$

$$\text{Rata-rata keterbacaan satu kalimat} = \frac{\text{Jumlah 7 skor keterbacaan}}{7}$$

Kemudian tiap kalimat dikelompokkan berdasarkan skor rata-rata yang diperoleh dari penilaian 7 rater, untuk tiap aspek. Dasar pengelompokkannya sama dengan yang dilakukan pada kelompok pengguna, yaitu dengan mengacu pada interval skor yang disajikan di sub-bab 3.6. Kalimat yang mendapatkan skor tertinggi dan

terendah menjadi contoh pembahasan secara kualitatif dengan tujuan menemukan alasan di balik skor tinggi/rendahnya itu.

Tahap terakhir adalah membandingkan hasil penilaian pengguna dengan hasil penilaian rater. Ini dilakukan dengan tujuan untuk mengetahui apakah pendapat mereka sama atau berbeda. Analisis dilakukan dengan melakukan uji statistika yang disajikan di sub-bab 3.7.

3.6 Penentuan Skala Interval

Untuk memudahkan penggolongan hasil penilaian pengguna terhadap masing-masing aspek penilaian terjemahan, maka dibuat interval skor yang mengacu pada nilai skala yang sudah dibahas di bagian 3.4.1. tabel 3.1 sampai dengan tabel 3.3.

$$\text{Interval} = \frac{H - L}{n_{\text{kategori}}} = \frac{5 - 1}{5} = 0,8$$

Amiruddin (2010); Riduwan & Akdo (2013)

Dengan demikian, selisih antara skala yang satu dengan yang lainnya adalah sebesar 0,8. Berdasarkan interval ini, maka tiap skor pada masing-masing skala memiliki interval sebagai berikut:

Tabel 3.2 Interval skor dalam tiap skala penilaian hasil terjemahan

Skala	Interval Skor	Keakuratan	Keberterimaan	Keterbacaan
5	4,21-5,00	Sangat akurat	Sangat berterima	Sangat mudah dipahami
4	3,41-4,20	Cukup akurat	Cukup berterima	Cukup mudah dipahami
3	2,61-3,40	Kurang akurat	Kurang berterima	Kurang dapat dipahami
2	1,81-2,60	Tidak akurat	Tidak berterima	Tidak dapat dipahami
1	1,00 -1,80	Sangat tidak akurat	Sangat tidak berterima	Sangat tidak dapat dipahami

3.7 Uji Statistika

Uji statistika dilakukan untuk dapat menjawab permasalahan nomor lima, yaitu dengan membandingkan penilaian pengguna dan rater, apakah terdapat perbedaan atau tidak. Perbandingan dilakukan pada dua aspek penilaian yaitu keberterimaan dan keterbacaan. Keakuratan tidak dibandingkan karena pengguna tidak ditanya akan aspek ini. Kelompok yang diuji di sini ada empat, yaitu tiga kelompok dari penilaian pengguna yang berasal dari bidang ilmu ekonomi, informatika, dan bahasa/sastra, dan satu lagi adalah kelompok penilaian rater. Satu alat untuk menguji kelompok pada kondisi ini adalah dengan menggunakan *anova one-way*.

Menurut Leech dkk. (2005, hlm. 129) *Anova* berasal dari *analysis of varian*. Mereka menyatakan bahwa untuk menguji perbedaan antara dua atau lebih kelompok (independen) dengan satu kelompok atau faktor dependen, maka uji *one-way anova* dapat dipakai. Dengan kata lain *one-way anova* merupakan salah satu uji perbandingan yang digunakan untuk menguji perbedaan *mean* (rata-rata) data lebih dari dua kelompok.

Maka dalam soal ini terdapat empat variabel independen (satu variabel dengan empat level), yaitu penilaian pengguna ekonomi, informatika, bahasa/sastra dan rater. Variabel dependennya adalah skor keberterimaan atau keterbacaan.

Dengan program SPSS, setelah menginput data dengan benar perbandingan tersebut dilakukan dengan cara:

*Analyze => Compare Means => One-Way ANOVA => pilih satu atau lebih variabel dependen dengan klik kiri lalu memindahkannya ke kotak *Dependent List* dengan cara mengklik panah sebelah atas di tengah kotak dialog => pilih *independent variable* dengan klik kiri lalu memindahkannya ke dalam kotak *Factor* (independent variable) dengan mengklik panah sebelah bawah di tengah kotak dialog => *Options* => pilih *Descriptives* dan *Homogeneity of variance test* => *Continue* => *OK*. (Leech dkk., 2005, hlm. 189)*

Hipotesa yang diajukan adalah sebagai berikut:

$$H_0 : \mu_R = \mu_E = \mu_K = \mu_B$$

$$H_1 : \mu_R \neq \mu_E \neq \mu_K \neq \mu_B$$

Di mana μ adalah mean atau rata-rata.

R = penilaian rater

E = penilaian pengguna ekonomi

K = penilaian pengguna informatika

B = penilaian pengguna bahasa/sastra

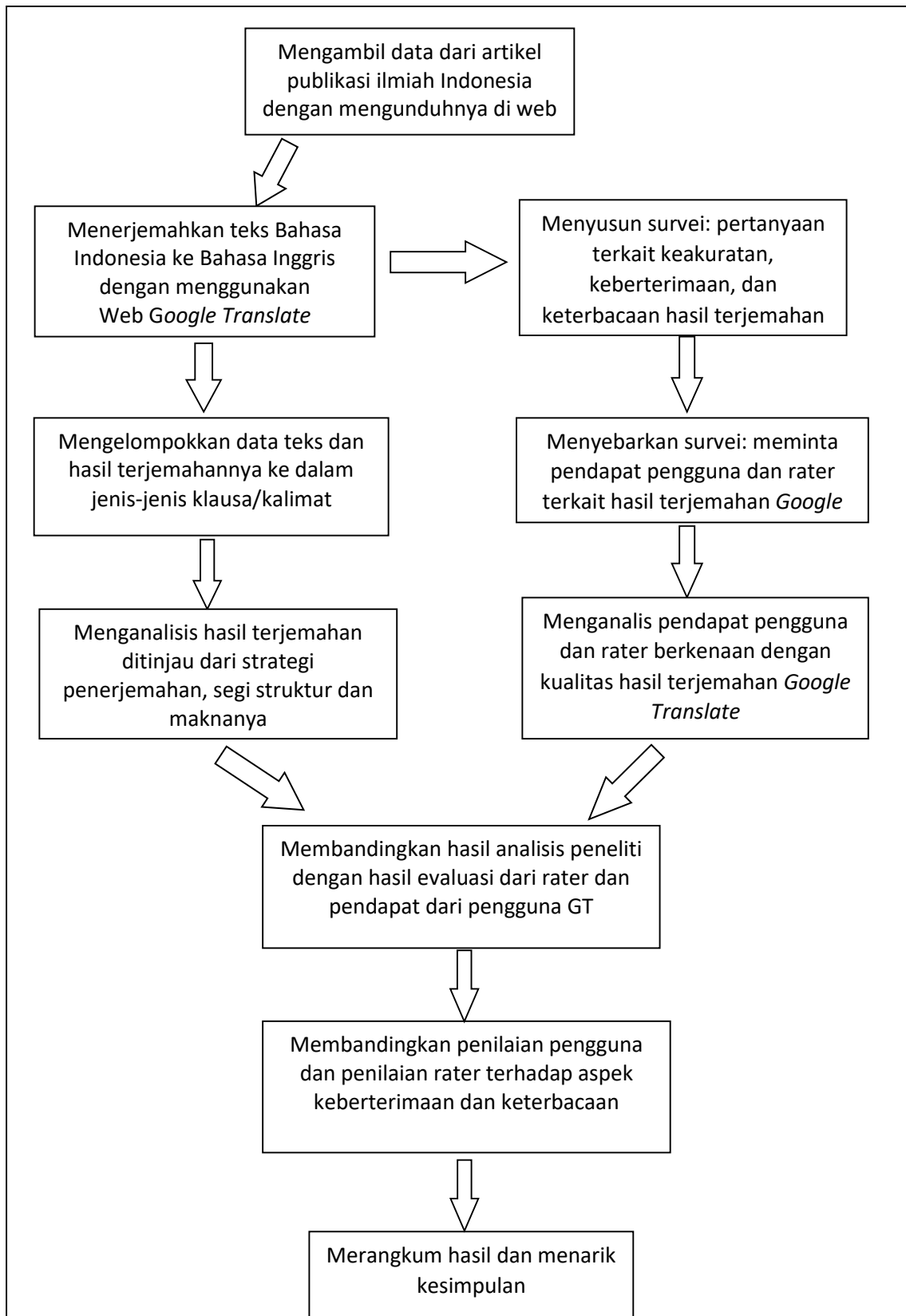
Hipotesa tersebut dapat dijabarkan sebagai berikut.

H₀: penilaian antara empat kelompok sama atau tidak ada perbedaan signifikan.

H₁: penilaian antara empat kelompok tidak sama atau berbeda secara signifikan

Bila ditetapkan nilai $p \leq 0,05$ dengan tingkat keyakinan 95%, maka bila dihasilkan nilai $p \leq 0,05$ maka H₀ harus diterima, artinya tidak terdapat perbedaan signifikan di antara penilaian empat kelompok. Sebaliknya, bila dihasilkan nilai $p > 0,05$ maka H₀ ditolak dengan makna bahwa terdapat perbedaan signifikan di antara penilaian empat kelompok. (Leech dkk., 2005, hlm. 53-54)

Dari uraian di bab III ini, keseluruhan tahap-tahap mulai dari proses pengumpulan data sampai analisis data dapat digambarkan dalam bagan berikut.



Gambar 3. 1 Langkah-langkah dalam prosedur penelitian