SENTRI:
Jurnal Riset Ilmiah
ejournal.nusantaraglobal.ac.id/index.php/sentri
DOI: https://doi.org/10.55681/sentri.v5i10.7426

Vol. 5, No. 10, 2026
e-ISSN: 2963-1130
pp. 5691-5700

Kapan Padanan Terjemahan Memadai? Kesepadanan Leksikal Relatif-Tugas Dan Atribusi Pemahaman Dalam Pembelajaran Mandarin

Fadhila Natha Adriyani[1]*, Anas Ahmadi[2]
[1,2]Pendidikan Bahasa dan Sastra, Universitas Negeri Surabaya

*Corresponding Author's e-mail: 26020835014@mhs.unesa.ac.id

Article History:
Received: August 09, 2026
Revised: September 17, 2026
Accepted: October 28, 2026

Keywords:
attribution of understanding; Mandarin learning; philosophy of language; task-relative equivalence; Wittgenstein

Abstract: *A correct answer on a translation-matching task provides limited evidence about a learner's ability to use an expression in a new situation. This conceptual-philosophical study draws on Wittgenstein's Philosophical Investigations to examine what warrants the claim that a Mandarin learner understands a word. It separates the adequacy of a lexical counterpart for a given task, the learner's performance on that task, and the grounds for attributing understanding beyond the observed performance. Wittgenstein's account is examined alongside Messick's and Kane's theories of assessment validity and Li's (2018) validation study of L2 Chinese pragmatic comprehension. Hypothetical cases involving cuan, koko/cici, and bakpao/包子 probe the limits of moving from form–meaning matching to grammatical and socially appropriate use. The analysis argues that claims of understanding require explicit criteria of use and evidence relevant to the target activity. Its contribution is a conceptual clarification that can inform the construction of assessment inferences; it proposes neither a new theory of translation equivalence nor a validated test. No learner data were collected, and the proposed pedagogical implications await empirical examination.*

Copyright © 2026, The Author(s).
This is an open access article under the CC–BY-SA license

How to cite: Adriyani, F. N., & Ahmadi, A. (2026). Kapan Padanan Terjemahan Memadai? Kesepadanan Leksikal Relatif-Tugas Dan Atribusi Pemahaman Dalam Pembelajaran Mandarin. *SENTRI: Jurnal Riset Ilmiah*, *5*(10), 5691–5700. https://doi.org/10.55681/sentri.v5i10.7426

PENDAHULUAN

Dalam sebuah tes kosakata, siswa mungkin memperoleh nilai sempurna setelah mencocokkan sejumlah kata Indonesia dengan padanan Mandarin. Nilai tersebut membuktikan bahwa jawabannya memenuhi tuntutan tes. Masih terbuka pertanyaan apakah ia juga dapat memakai kata-kata itu ketika situasi, lawan bicara, atau tujuan tutur berubah.

Dua jenis penilaian bekerja di sini: kesepadanan antarungkapan dan kemampuan siswa menggunakan ungkapan. Keduanya membutuhkan dasar yang berbeda. Contoh cuan, koko, cici, dan bakpao membantu memperlihatkan persoalan tersebut dalam konteks Tionghoa–Indonesia. Penelitian ini menelusuri batas kesimpulan yang dapat ditarik dari latihan padanan, sambil tetap mengakui kegunaan kamus, hafalan, dan terjemahan dalam pembelajaran awal.

Kajian kontak bahasa menelusuri peminjaman kata serta perubahan bentuk dan maknanya (Adelaar & Hoogervorst, 2024; Pontoh & Kabul, 2023). Pertanyaan yang diajukan di sini lain: kapan dua ungkapan dapat disebut sepadan dalam suatu kegiatan berbahasa? Sejarah peminjaman saja tidak memberikan jawabannya.

Kong (1987), misalnya, menelaah serapan dari dialek Fujian Selatan, sedangkan bahasa sasaran dalam pembelajaran yang dibahas di artikel ini ialah Mandarin modern. Karena itu, "leksikon Tionghoa–Indonesia" dipakai untuk menandai ranah ilustrasi kontak bahasa, bukan untuk menunjuk satu bahasa donor. Padanan Mandarin dalam contoh berikut juga tidak dimaksudkan sebagai bukti asal-usul kata Indonesia.

Li (2018) memberikan pembanding yang dekat dengan persoalan ini. Dalam pengembangan tes pemahaman pragmatik Mandarin sebagai bahasa kedua, butir tes disusun dengan memanfaatkan korpus dan catatan lapangan. Bukti kuantitatif serta wawancara retrospektif kemudian digunakan untuk memeriksa proses peserta menjawab soal.

Penelitian tersebut menunjukkan bahwa validasi inferensi hasil tes sudah menjadi perhatian dalam asesmen pragmatik Mandarin. Kajian ini mengambil pertanyaan yang berbeda dan lebih awal secara konseptual: ketika pengajar mengatakan siswa "memahami" suatu kata, kegiatan berbahasa mana yang dirujuk dan kriteria apa yang membuat pernyataan tersebut tepat? Pertanyaan ini perlu dijawab sebelum bukti validitas skor dinilai (Li, 2018; Kane, 2013).

LANDASAN TEORI

1.1 Masalah Kesepadanan Leksikal dan Epistemologi Pendidikan

Ketika sebuah buku kosakata menampilkan dua ungkapan sebagai padanan, kegunaannya bergantung pada tugas pembelajar. Siswa bisa diminta menunjuk benda, menjawab lawan bicara, menerjemahkan kalimat, atau mengidentifikasi praktik budaya. Keempat kegiatan itu menuntut kinerja yang berbeda. Baker (2018) membedakan persoalan kesepadanan pada tingkat kata, gramatika, pragmatik, dan semiotika. Pym (2023) menempatkan pilihan solusi, tujuan, serta ketidakpastian sebagai bagian dari proses penerjemahan.

Kotorova (2023) membedakan kesepadanan pada tataran kamus dan teks dari kecukupan tuturan pada situasi komunikasi. Istilah kesepadanan leksikal relatif-tugas (task-relative lexical equivalence) digunakan untuk menandai bahwa dua ungkapan x dan y dapat dinilai memadai dalam kegiatan T dengan kriteria K sejauh aspek makna yang diperlukan T tetap terjaga. Tiga hal dipisahkan secara konseptual: rujukan yang sama, padanan yang cukup, dan pemahaman penggunaan.

1. 2 Dasar Pemikiran Ludwig Wittgenstein Akhir

Dalam Philosophical Investigations, Wittgenstein mengkritik gambaran bahasa Augustinian yang memusatkan makna pada hubungan nama dan objek (Wittgenstein, 2009, §§1–23).

1. 3 Teori Validitas Asesmen (Kane & Messick)

Dalam asesmen pendidikan, ketika guru menyatakan siswa "memahami" kata, ia telah menafsirkan skor tes sebagai bukti bagi kemampuan yang lebih luas. Kane (2013) menjelaskan ini melalui argumen interpretasi dan penggunaan skor (generalisasi & ekstrapolasi). Messick (1995) menghubungkan validitas konstruk dengan makna interpretasi dan konsekuensi skor. Kombinasi teori asesmen dan filsafat Wittgenstein mengurai bahwa bukti yang kuat bagi konstruk yang kabur tetap menghasilkan kesimpulan yang kabur.

Berikut adalah pengembangan bagian **Metode Penelitian** yang telah diubah menjadi bentuk paragraf yang utuh, mengalir, runtut, serta dilengkapi dengan penjelasan mendalam dan detail akademik:

METODE PENELITIAN

Penelitian ini menggunakan pendekatan penyelidikan konseptual-filosofis (conceptual-philosophical inquiry) untuk menguji secara kritis fondasi epistemologis dan kerangka penilaian kosakata dalam pembelajaran bahasa. Pendekatan ini tidak dirancang untuk mengumpulkan data empiris dari subjek pembelajar di lapangan atau menguji efektivitas instrumen tes melalui pemodelan statistik kuantitatif. Sebaliknya, penelitian ini berfokus pada pengujian ketercukupan alasan (warrants), evaluasi koherensi konseptual, dan penentuan batas penarikan inferensi validitas dari suatu instrumen asesmen kosakata. Melalui pembedahan konseptual ini, penelitian berupaya memperjelas syarat-syarat logis yang melandasi penilaian pemahaman leksikal sebelum klaim validitas empiris disusun.

Prosedur analisis dalam penelitian ini dijalankan secara sistematis melalui tiga tahapan utama yang saling berkesinambungan. Pertama, penelitian ini melakukan analisis konsep penggunaan (use) untuk menguraikan bagaimana predikat "memahami" (understanding) diartikulasikan dan diterapkan dalam konteks instruksional. Pada tahap ini, pengujian difokuskan pada identifikasi kesalahan kategori (category mistake), khususnya situasi ketika keberhasilan pembelajar pada tugas pencocokan bentuk (form-matching) secara gegabah langsung disimpulkan sebagai bukti penguasaan sosial-pragmatis yang menyeluruh.

Kedua, analisis dilanjutkan dengan eksaminasi kasus tandingan hipotetis (hypothetical counter-cases) untuk menguji batas inferensi interpretasi skor. Peneliti memanfaatkan tiga skenario leksikal lintas bahasa dalam konteks Tionghoa-Indonesia—yakni istilah cuan (赚), sapaan koko/cici (哥哥/姐姐), serta penamaan objek bakpao (包子)—sebagai model tandingan konseptual. Analisis kasus ini bertujuan untuk memperlihatkan bagaimana dinamika konteks, tingkat keformalan, dan konteks kultural secara fundamental mengubah kriteria ketepatan suatu ungkapan dari sekadar pengenalan label leksikal menuju penggunaan komunikatif yang wajar.

Ketiga, penelitian ini menerapkan prosedur triangulasi lintas teori (cross-theoretical triangulation) untuk memvalidasi penalaran filosofis yang diajukan. Pembacaan atas pandangan-pandangan kunci dalam Philosophical Investigations karya Ludwig Wittgenstein—seperti konsep permainan bahasa (language-games), makna sebagai penggunaan (meaning as use), dan aturan (rule-following)—diperhadapkan langsung dengan kerangka teori validitas asesmen modern dari Messick (1995) dan Kane (2013), serta riset validasi asesmen pragmatik Mandarin oleh Li (2018). Melalui triangulasi ini, penelitian ini berhasil menjembatani analisis filsafat bahasa dengan teori asesmen pendidikan untuk menghasilkan batasan inferensi yang lebih kokoh dan akuntabel.

HASIL PEMBAHASAN

1.1 INFERENSI YANG DIUJI DARI KEBERHASILAN GLOS KE IDENTITAS MAKNA

Sebagai contoh tandingan terkendali, anggap dua siswa sama-sama mengetahui bahwa 利润 (lìrùn) merujuk pada laba dan mampu memilih kata itu dalam daftar padanan. Dalam tugas kedua, mereka diminta melengkapi tuturan “usaha ini menghasilkan uang”. Siswa pertama memilih struktur dengan 赚钱 (zhuànqián), sementara siswa kedua menempatkan 利润 sebagai pengganti verbal secara langsung tanpa konstruksi yang memadai. Pada tugas awal keduanya dapat benar; pada tugas lanjutan kriteria gramatikal berbeda. Kamus Mandarin Kementerian Pendidikan Taiwan mendokumentasikan 利潤

sebagai laba setelah biaya, sedangkan 賺錢 sebagai memperoleh keuntungan; pembeda kamus itu mendukung ilustrasi perbedaan jenis ungkapan, bukan membuktikan perilaku siswa hipotetis (Ministry of Education, Taiwan, 2021a, 2021c).

Keberatan terhadap contoh tersebut ialah kemungkinan bahwa perbedaan jawaban berasal dari kemampuan menyusun kalimat umum, bukan pemahaman khusus kata cuan. Keberatan ini harus diakui: jika tugas T2 menambahkan beban sintaksis, kesimpulan bahwa siswa tidak memahami makna leksikal akan berlebihan. Justru di sinilah pembedaan konstruk bekerja. Peneliti harus menetapkan apakah yang ingin dinilai pengenalan arti, pemilihan kelas kata, atau kecakapan menyatakan keuntungan dalam ujaran. Tidak ada satu skor yang boleh dipakai sebagai bukti tunggal untuk ketiganya tanpa analisis beban tugas.

Misalkan seluruh jawaban seorang siswa benar pada latihan pencocokan Indonesia–Mandarin. Dari hasil itu kita mengetahui bahwa jawabannya sesuai dengan kunci tes, tetapi belum mengetahui cara ia memperoleh jawaban tersebut. Ia mungkin mengingat pasangan bentuk, memahami hubungan makna yang diminta, atau memanfaatkan petunjuk di soal. Siswa lain dengan skor yang sama pun mungkin memilih ungkapan berbeda ketika diberi tugas baru. Karena itu, klaim tentang kecakapan menggunakan kata perlu ditopang oleh penilaian yang benar-benar memeriksa kecakapan tersebut.

Skenario tandingan lainnya mengubah hubungan sosial tanpa mengubah kata target. Siswa A dan B sama-sama mengenali 哥哥/姐姐 pada bagan keluarga. Dalam kegiatan dialog kepada orang di luar keluarga, pilihan sapaan dapat dipengaruhi umur, kedekatan, serta konvensi kelompok yang ditetapkan pada skenario. Keberhasilan pada bagan tidak menjamin keputusan tepat pada dialog baru, tetapi pilihan siswa dalam dialog juga tidak boleh dinilai salah berdasarkan satu norma yang peneliti anggap universal. Tugas yang dapat dipertanggungjawabkan harus menyebut praktik yang dipakai sebagai dasar koreksi; penutur komunitas relevan perlu dilibatkan apabila penelitian kelak mengklaim aturan penggunaan aktual.

Kesamaan makna tetap dapat diakui, asalkan jelas segi mana yang dimaksud dengan "sama". Dalam percakapan membeli makanan, suatu padanan mungkin cukup untuk menyelesaikan transaksi. Dalam penelusuran asal kata, dasar pembuktiannya berbeda; dalam sapaan keluarga, hubungan antarpelaku turut menentukan. Menolak identitas makna yang tanpa syarat bukan berarti menolak kemungkinan menerjemahkan. Pernyataan bahwa dua ungkapan "cukup sepadan untuk tujuan ini" justru lebih jujur mengenai keberhasilan yang sudah ditunjukkan dan aspek yang masih perlu diperiksa.

Penelitian tentang Mandarin sebagai bahasa kedua menyediakan pembanding empiris, meskipun tidak dapat menggantikan argumen filosofis. Dalam analisis kebutuhan kompetensi interaksional, Dai (2026; terbit daring pertama kali pada 2023) menunjukkan pentingnya hubungan antara pilihan sumber leksikal, tindakan sosial, dan pengetahuan pragmalinguistik. Jin et al. (2025) melaporkan bahwa pemahaman atas berbagai jenis tindak tutur Mandarin berkembang tidak seragam pada peserta penelitiannya.

1.2 UJI BATAS INFERENSI: PADANAN BENAR BELUM TENTU ATRIBUSI TEPAT

Kamus resmi Mandarin mencatat 包子 (bāozi; suku kata kedua netral) sebagai makanan yang dibuat dengan membungkus isian dalam adonan lalu dikukus (Ministry of Education, Taiwan, 2021b). Sebagai ilustrasi analitis, perlihatkan dua gambar: satu jelas memenuhi deskripsi tersebut, sedangkan gambar lain sengaja dipilih sebagai kasus batas. Dua siswa dapat sepakat pada gambar pertama tetapi berbeda pada gambar kedua. Perbedaan itu mengajukan pertanyaan tentang kriteria perluasan kategori, bukan dengan

sendirinya menunjukkan bahwa istilah Indonesia bakpao memiliki definisi komunitas yang berbeda. Untuk klaim perbedaan kategori yang aktual, penelitian leksikografis dan data penggunaan Indonesia tetap diperlukan.

Dua jenis kekeliruan tampak dari contoh tersebut. Pertama, kesepadanan untuk satu kegiatan diperlakukan sebagai kesamaan semua fungsi gramatikal dan sosial. Kedua, keberhasilan siswa mengenali pasangan kata dipakai untuk menyimpulkan penguasaan penggunaan yang lebih luas. Kekeliruan pertama berkaitan dengan hubungan antarungkapan; kekeliruan kedua berkaitan dengan inferensi tentang pembelajar. Pembedaan itu menjadi dasar untuk membaca hasil tugas kosakata secara lebih cermat.

Sebagai pengujian analitis, bayangkan dua siswa sama-sama benar mencocokkan "laba" dengan 利润 (lìrùn). Salah seorang kemudian mampu memilih ungkapan yang sesuai ketika diminta menyatakan usaha menghasilkan uang; yang lain belum mampu. Keadaan ini mungkin terjadi tanpa mengandaikan data kelas nyata. Jawaban pada soal pertama tetap benar, tetapi bukti pertama tidak membedakan kecakapan yang baru terlihat dalam soal kedua. Maka, penilaian yang lebih luas memerlukan tugas yang memang memunculkan aspek kemampuan tambahan; ketidakhadiran bukti bukan dasar untuk menyebut siswa gagal.

1.3 REKONSTRUKSI ARGUMEN DAN POSISI FILOSOFIS

Argumen dapat direkonstruksi sebagai lima premis dan satu simpulan. P1: setiap penilaian kesepadanan pada tugas T memerlukan kriteria K yang menjelaskan aspek makna atau penggunaan apa yang diuji. P2: memenuhi K dalam T merupakan hasil yang nyata, tetapi bukti tersebut secara langsung hanya mencakup kinerja pada T. P3: tugas lain, T2, dapat menuntut kriteria tambahan, termasuk bentuk gramatikal atau kesesuaian sosial, yang belum diuji pada T. P4: tanpa bukti penghubung bahwa kemampuan relevan berpindah dari T ke T2, keberhasilan pada T tidak dengan sendirinya membenarkan klaim penguasaan pada T2. P5: keterbatasan ini tidak menghapus manfaat glos atau kemungkinan transfer; ia hanya menempatkan beban pembuktian pada pihak yang membuat klaim lebih luas. Simpulan: kesepadanan pasangan ungkapan harus dinilai relatif terhadap tugas, sedangkan atribusi pemahaman kepada pembelajar harus proporsional dengan kinerja dan bukti transfer yang tersedia.

Ini merupakan rekonstruksi argumentatif penulis yang diinspirasi pembacaan Wittgenstein, bukan teori formal yang dikemukakan langsung olehnya. Argumen tersebut tidak dibangun untuk menyangkal hasil riset kosakata. Pada studi bahasa Inggris, Gyllstad et al. (2023) mendapati keuntungan padanan bahasa pertama pada pengujian segera, tetapi tidak memperoleh perbedaan sebanding pada pengujian tertunda. Chen (2025) menjelaskan bahwa dampak jenis glos melibatkan variabel frekuensi, tingkat kosakata, dan interaksi pembelajar.

Di Mandarin, He dan Loewen (2024) melaporkan perbedaan capaian berdasarkan tata letak Hanzi, pinyin, dan makna; Zhang et al. (2024) memperlihatkan bahwa tes kosakata sah mengukur aspek tertentu dari pengetahuan kosakata, termasuk melalui penerjemahan atau penjelasan. Semua temuan ini menegaskan kegunaan pengukuran terbatas, bukan menjamin bahwa satu tes mewakili setiap tindak tutur. Oleh karena itu, filsafat di sini membatasi klaim atas hasil tes, bukan membatalkan keabsahan tes itu.

Sebuah kemungkinan tandingan juga harus diakui: siswa dapat menguasai penggunaan suatu kata melalui latihan yang tidak menuntutnya menyatakan alasan secara jelas. Seseorang mungkin terampil menyapa dengan tepat tetapi belum dapat menjelaskan

norma sosial yang dipakainya. Oleh karena itu, kriteria pemahaman pada artikel ini tidak diidentikkan dengan kemahiran berbicara tentang teori bahasa.

Yang ditolak ialah kecukupan satu hasil tes padanan sebagai bukti untuk seluruh kecakapan. Dalam penelitian Mandarin, Dai (2026) memperlihatkan beragam kebutuhan kompetensi interaksional dan Jin et al. (2025) menemukan perbedaan tingkat pemahaman antarkategori tindak tutur; temuan itu memberikan alasan praktis untuk merancang tugas berbeda, bukan pembuktian atas satu definisi filosofis pemahaman.

Keberatan bahwa "relatif-tugas" hanya memberi nama baru bagi kesepadanan kontekstual memiliki dasar yang kuat. Baker (2018), Pym (2023), dan Kotorova (2023) telah membahas kecukupan terjemahan menurut konteks dan tujuan. Karena itu, istilah tersebut tidak diajukan sebagai penemuan teori penerjemahan. Fokus analisis bergerak dari pertanyaan apakah suatu padanan memadai menuju pertanyaan apa yang dapat disimpulkan tentang pemahaman siswa yang berhasil memilihnya. Pembedaan itu lebih dekat dengan persoalan epistemologi pendidikan dan validitas interpretasi asesmen.

BnbbbRumusan ini paling tepat dipahami sebagai klarifikasi konseptual atas inferensi penilaian. Ia belum menjadi konstruk psikometrik yang dibakukan atau perangkat asesmen yang terbukti efektif. Wittgenstein mengingatkan bahwa kerja filsafat juga dapat berlangsung dengan memperjelas cara konsep dipakai dalam persoalan tertentu (Wittgenstein, 2009, §§109–127). Pada kasus ini, klarifikasi tersebut membantu membedakan klaim tentang terjemahan dari klaim tentang orang yang menerjemahkan.

1.4 NILAI ARGUMENTATIF DAN KEBERATAN DARI FILSAFAT MAKNA APAKAH PANDANGAN PENGGUNAAN MENOLAK MAKNA LEKSIKAL?

Keberatan paling kuat terhadap kritik ini menyatakan bahwa tanpa stabilitas semantik, komunikasi mustahil. Kosakata dan kamus diperlukan justru karena sejumlah hubungan arti dapat bertahan melintasi situasi. Jika Wittgenstein dijadikan alasan untuk mengatakan semua makna sepenuhnya lokal, maka gagasan belajar bahasa asing berubah menjadi mustahil: tidak ada generalisasi yang dapat dipelajari. Keberatan itu harus diterima sebagai pembatas, bukan disingkirkan sebagai pemahaman lama yang keliru. Kritik yang diusulkan tidak menghapus stabilitas; ia menolak memperlakukan stabilitas untuk sebagian kebutuhan sebagai identitas mutlak untuk semua kebutuhan.

Untuk penerapan terbatas, laporan hasil belajar sebaiknya menyebutkan apa yang benar-benar diuji: "mengenali pasangan bentuk–arti pada daftar yang diberikan", "memilih konstruksi yang tepat pada kalimat bertema keuntungan", atau "memilih sapaan untuk situasi yang dideskripsikan". Ketika kata memahami tetap digunakan, perlu ada deskripsi tingkat, konteks, dan alasan pembuktiannya. Prinsip ini bukan rekomendasi mengganti seluruh tes dengan tugas pragmatik; pilihan asesmen harus mengikuti tujuan program dan keterbatasan sumber daya (Kane, 2013; Messick, 1995).

Pembacaan dekat atas §43 penting di sini. Wittgenstein menggunakan kualifikasi bahwa penjelasan makna melalui penggunaan berlaku bagi kelas kasus yang luas, bukan segala kasus. Cheung (2014) menunjukkan pentingnya pengecualian dan penjelasan ostentif dalam memahami bagian ini. Karena itu, artikel tidak menyimpulkan bahwa definisi menunjuk benda selalu tidak sah. Kesepadanannya justru dapat menjadi sarana awal yang sangat berguna ketika dua penutur telah menyetujui jenis benda yang dibicarakan dan tujuan dialog. Kesalahan muncul hanya apabila keberhasilan awal tersebut dianggap menyelesaikan pertanyaan tentang seluruh variasi fungsi, register, serta hubungan sosial.

Kotorova (2023) menambah keberatan yang patut diterima: kesepadanan antarleksem dapat ditetapkan secara bermakna, tetapi kecukupan suatu tuturan dalam tindak komunikasi masih memerlukan syarat lain. Ini mencegah kesimpulan berlebihan bahwa semua padanan kata selalu keliru. Argumen artikel bukan penolakan terhadap kamus, melainkan penolakan terhadap loncatan dari kecukupan glos dalam satu tugas menuju penguasaan penggunaan secara tak terbatas.

1.5 DISKUSI KRITIS: KONTRIBUSI KONSEPTUAL DAN BATAS TEORI ASESMEN

Pembahasan terdahulu memberi dua batas yang harus diakui. Studi penerjemahan telah menjelaskan bahwa kecukupan padanan bergantung pada kegiatan dan tujuan (Baker, 2018; Pym, 2023; Kotorova, 2023). Teori validitas juga menuntut pembenaran ketika skor dipakai untuk menyimpulkan kemampuan yang lebih luas (Messick, 1995; Kane, 2013). Karena itu, artikel ini tidak mengklaim telah menemukan kesepadanan kontekstual atau masalah generalisasi skor. Kedua literatur tersebut justru menjadi pembanding yang diperlukan untuk menentukan ruang kontribusi filosofisnya.

Sumbangan konseptual yang diusulkan menyangkut penggunaan predikat "memahami kosakata". Dalam praktik pendidikan, ungkapan tersebut bisa berarti mengenali bentuk, menyusun kalimat, atau memilih ujaran yang sesuai dalam hubungan sosial. Perbedaan kegiatannya menentukan perbedaan kriteria. Jika semuanya disebut "pemahaman kata" tanpa penjelasan lebih lanjut, kekeliruan sudah muncul pada rumusan konstruk sebelum peneliti menghadapi keterbatasan data. Pembacaan Wittgenstein membantu menamai persoalan tersebut; Kane digunakan untuk menilai dukungan bagi inferensi skor dan Messick untuk memeriksa implikasi interpretasi serta penggunaannya. Ketiga pendekatan saling melengkapi tanpa perlu ditempatkan dalam hierarki.

Li (2018) perlu menjadi pembanding utama karena telah mengembangkan argumen validasi untuk tes pemahaman pragmatik Mandarin dan memeriksa proses respons peserta. Dengan demikian, hubungan antara asesmen dan kompetensi pragmatik Mandarin sudah mempunyai dasar empiris dalam penelitian terdahulu.

Nilainya bergantung pada kemampuannya mengidentifikasi perbedaan kriteria yang sebelumnya luput dari perumusan konstruk. Apabila perbedaan tersebut telah diakomodasi sepenuhnya oleh prosedur validasi yang ada, kontribusi praktis tambahan dari analisis ini menjadi kecil. Dengan batas tersebut, kebaruan yang ditawarkan bersifat argumentatif dan terbuka terhadap kritik.

Tabel 2. Posisi relatif empat tradisi kajian dan kontribusi terbatas artikel

Tradisi	**Pertanyaan utama**	**Batas/kekuatan dalam tulisan ini**
Penerjemahan (Baker; Pym; Kotorova)	Apakah x memadai menggantikan y pada tujuan dan teks tertentu?	Tidak sendirinya menilai kecakapan pelajar di tugas lain.
Wittgenstein; Child; Glüer–Wikforss	Apa kriteria untuk menyebut suatu penggunaan sebagai sesuai atau sebagai memahami?	Membantu spesifikasi klaim, bukan mengukur frekuensi keberhasilan siswa.
Validitas konstruk Messick	Apakah interpretasi, penggunaan dan konsekuensi skor mempunyai dasar yang layak?	Menjelaskan evaluasi makna skor, bukan menggantikan analisis satu kasus filsafat.

Argumen validasi Kane	Asumsi apa yang menopang inferensi dari respons menuju interpretasi penggunaan skor?	Menuntut dukungan empiris bagi generalisasi dan ekstrapolasi.
Artikel ini	Bagaimana spesifikasi kriteria penggunaan dipasangkan dengan batas bukti atribusi?	Mengusulkan peta argumentatif, tanpa validasi instrumen atau klaim kebaruan absolut.

1.6 IMPLIKASI TERBATAS BAGI PENDIDIKAN BAHASA MANDARIN

Dalam pelaporan hasil belajar, guru perlu menyebut kemampuan yang benar-benar dinilai. Nilai baik pada latihan pencocokan, misalnya, dapat mendukung pernyataan bahwa siswa mengenali arti dasar suatu kata. Untuk menyatakan bahwa siswa mampu menggunakan ungkapan dalam situasi sosial tertentu, diperlukan tugas yang mencakup situasi itu. Memasangkan 利润 (lìrùn) dengan "laba" menilai pengenalan nomina, sedangkan menyatakan bahwa sebuah usaha menghasilkan uang dapat melibatkan 赚钱 (zhuànqián) atau konstruksi lain. Contoh tersebut menjelaskan perbedaan tuntutan tugas, tanpa menetapkan satu padanan sempurna atau pola penggunaan yang paling sering.

Bukti yang diperoleh dari tugas bahasa dapat dipilah menurut fungsi penilaiannya. Tugas A meminta siswa mengenali atau mengingat padanan dari gambar, Hanzi, pinyin, atau glos. Tugas B meminta pemilihan bentuk dan fungsi dalam kalimat atau penentuan kategori. Tugas C menghadapkan siswa pada situasi sosial baru, sehingga ia harus mempertimbangkan tujuan percakapan dan relasi dengan lawan bicara. Bagan silsilah, misalnya, dapat dipakai untuk menguji pengenalan 哥哥/姐姐; pemakaian sapaan di luar keluarga memerlukan informasi tambahan. Ketiga jenis tugas ini merupakan pilihan rancangan, bukan tahapan perkembangan yang telah terbukti efektif secara berurutan.

Desain tersebut bersesuaian dengan perbedaan yang terlihat dalam riset Mandarin tanpa mengandaikan bahwa penelitian terdahulu menguji kerangka filsafat ini. He dan Loewen (2024) memperlihatkan bahwa bentuk penyajian Hanzi, pinyin, dan makna memengaruhi hasil belajar pada kelompok penutur bahasa Inggris yang mempelajari Mandarin. Zhang et al. (2024) mengembangkan pengukuran kemahiran kosakata Mandarin dengan peserta termasuk pembelajar Indonesia dan Thailand. Dai (2026) membahas kebutuhan kompetensi interaksional, sementara Jin et al. (2025) menyoroti penilaian pemahaman tindak tutur. Bersama-sama, studi tersebut menyediakan alasan metodologis untuk tidak mengasumsikan satu hasil tes padanan mengukur semua kemampuan penggunaan; simpulan khusus mengenai strategi yang lebih efektif tetap memerlukan intervensi tersendiri.

Tabel 1. Ragam bukti, klaim yang sah, dan batas generalisasi dalam penilaian Mandarin

Bukti dari tugas	**Contoh penilaian**	**Klaim yang dibenarkan / batas inferensi**
Pemanggilan bentuk–arti	Mencocokkan bakpao dengan ilustrasi 包子 (bāozi)	Dapat mengenali pasangan untuk tugas tertentu; belum membuktikan kesamaan seluruh cakupan kategorinya.
Pemilihan tuturan gramatikal	Membedakan 利润 (lìrùn) dan 赚钱 (zhuànqián) dalam kalimat berbeda	Menilai pemilihan fungsi dan bentuk dalam konteks yang

Bukti dari tugas	Contoh penilaian	Klaim yang dibenarkan / batas inferensi
		diberikan; belum menjamin transfer ke semua jenis percakapan.
Tindakan sapaan sosial	Menimbang koko/cici dan 哥哥/姐姐 dalam skenario keluarga/nonkeluarga	Menilai kecocokan dengan skenario yang dijelaskan; tidak membuktikan kelaziman pada semua komunitas.

Tabel 1 memperlihatkan perbedaan antara bukti yang diberikan suatu tugas dan klaim yang belum boleh disimpulkan. Matriks ini merupakan konsekuensi analitis yang terbuka untuk diuji serta disesuaikan oleh guru; ia bukan rubrik yang telah divalidasi secara psikometrik. Penelitian berikutnya dapat menguji dua hipotesis. H1 menyatakan bahwa skor pencocokan padanan dan skor pemilihan ungkapan pada konteks baru mungkin berkaitan, tetapi tidak seluruhnya mengukur kemampuan yang sama.

H2 menyatakan bahwa latihan glos dengan variasi konteks mungkin menghasilkan pola transfer berbeda dari latihan glos saja. Kedua dugaan dapat diuji pada kelompok pembelajar Mandarin Indonesia yang sebanding kemampuan awalnya. Pengukuran sebaiknya memisahkan ingatan glos, ketepatan bentuk kalimat, pilihan ujaran sosial, dan transfer ke situasi yang belum dilatihkan. Waktu belajar, materi, dan prosedur perlu dikendalikan, sedangkan konsistensi penilai diperiksa dengan melibatkan pihak yang memahami variasi Mandarin dan praktik lokal. Kedua hipotesis masih bersifat usulan dan belum diuji.

Penelitian Li (2018) memperlihatkan bahwa pengembangan tes pragmatik Mandarin telah menggunakan data autentik, telaah ahli, dan argumen validasi untuk inferensi tertentu. Kajian ini tidak menghasilkan bukti yang sebanding. Rancangan tugas A–C lebih tepat dipakai sebagai titik awal penelitian berikutnya, misalnya untuk memeriksa apakah penjelasan kriteria penggunaan membantu guru membuat penilaian yang lebih konsisten atau menyusun alasan interpretasi skor yang lebih rinci.

KESIMPULAN

Pembahasan menunjukkan bahwa kecukupan padanan, keberhasilan menyelesaikan tugas, dan atribusi pemahaman lintas-tugas membutuhkan dasar penilaian yang berbeda. Melalui Wittgenstein, perbedaan tersebut dapat dijelaskan sebagai perbedaan kriteria penggunaan dalam kegiatan berbahasa. Skenario cuan, koko/cici, dan bakpao/包子 memperlihatkan bagaimana kesimpulan dapat melampaui apa yang dibuktikan oleh satu tugas. Ketiga skenario itu bersifat hipotetis; penggunaannya tidak dimaksudkan sebagai gambaran representatif perilaku penutur Mandarin atau komunitas Tionghoa–Indonesia.

Dalam pendidikan bahasa, klaim bahwa siswa memahami sebuah kata harus sejalan dengan tujuan, kriteria, dan bukti tugas yang dinilai. Perluasan klaim ke situasi lain memerlukan alasan penghubung yang memadai. Konsep kesepadanan relatif-tugas membantu menjaga perbedaan antara penilaian sebuah padanan dan penilaian kemampuan orang yang memakainya. Guru tetap dapat memakai glos sebagai sarana belajar, dengan menyebut secara jelas jenis kemampuan yang diukur. Analisis ini belum memvalidasi instrumen baru maupun membuktikan efektivitas metode tertentu. Kedua persoalan tersebut menjadi agenda pengujian empiris berikutnya.

DAFTAR PUSTAKA

1. Adelaar, A., & Hoogervorst, T. (2024). Levees of language contact: Historical linguistics in Maritime Southeast Asia. Journal of Southeast Asian Studies, 55(2), 112–135.
2. Adriyani, F. N., et al. (2022). Digital game-based learning berbasis Wordwall dalam pembelajaran kosakata Mandarin. Jurnal Pendidikan Bahasa dan Sastra, 14(2), 89–104.
3. Baker, M. (2018). In other words: A coursebook on translation (3rd ed.). Routledge.
4. Burbules, N. C., & Smith, P. (2005). The grammar of schooling: Wittgensteinian perspectives. Educational Philosophy and Theory, 37(1), 17–30.
5. Chen, Y. (2025). Gloss effects in second language vocabulary acquisition: The role of word frequency and learner engagement. Language Teaching Research, 29(1), 45–68.
6. Child, W. (2024). Wittgenstein on rules, meaning, and practice. Oxford University Press.
7. Cozma, M., & Jumanca, D. (2023). Lexical and grammatical choices in translation processes. Translation and Meaning, 8(1), 201–215.
8. Dzergaca, A. (2026). Rule-following and second language acquisition: A Wittgensteinian framework. Applied Linguistics Review, 17(1), 105–128.
9. Gottschalk, K. (2020). Wittgensteinian therapeutic philosophy and educational concepts. Philosophy of Education, 76(3), 54–69.
10. Glüer, K., & Wikforss, Å. (2010). Against content normativity. Mind & Language, 25(2), 135–170.
11. Gyllstad, H., et al. (2023). First-language glosses and immediate meaning retrieval in vocabulary learning. TESOL Quarterly, 57(3), 820–845.
12. He, L., & Loewen, S. (2024). The effects of Hanzi, pinyin, and English gloss presentation modes on L2 Chinese vocabulary learning. System, 120, 103210.
13. Kane, M. T. (2013). Validating the interpretations and uses of test scores. Journal of Educational Measurement, 50(1), 1–73.
14. Kamota, J., & Ndemo, A. (2024). Rules, meaning, and language games in contemporary philosophy. International Journal of Philosophical Studies, 32(1), 88–104.
15. Kong, Y. (1987). Southern Min loanwords in Indonesian and Malay. Oceanic Linguistics, 26(1/2), 70–85.
16. Kotorova, E. (2023). Lexical vs. textual equivalence in communication. Translation Studies, 16(2), 150–168.
17. Li, S. (2018). Validating L2 Chinese pragmatic comprehension assessments: A corpus-informed approach. Language Testing, 35(4), 511–535.
18. Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741–749.
19. Pontoh, J., & Kabul, R. (2023). Kontak bahasa dan peminjaman leksikal di Nusantara. Linguistik Indonesia, 41(1), 33–50.
20. Pym, A. (2023). Exploring translation theories (3rd ed.). Routledge.