

## 'Huh?' pada Mesin: Strategi Perbaikan Dialog (Repair) di Voice Assistant Lintas Bahasa



Check for updates

Randi Perkasa <sup>a,1\*</sup>, Yudha Putra Indarto <sup>a</sup>, Kalmiatin Musyawari <sup>a</sup>

<sup>a</sup> Program Studi Desain Komunikasi Visual, Universitas Komputer Indonesia, Indonesia

<sup>1</sup> [randiperkasa@gmail.com](mailto:randiperkasa@gmail.com)\*

\* Corresponding Author

### ABSTRACT

Penelitian ini mengeksplorasi penggunaan strategi perbaikan dialog (repair strategies) pada asisten suara lintas bahasa, dengan fokus pada interaksi pengguna yang mengekspresikan kebingungannya melalui respons "Huh?". Sebagai alat komunikasi berbasis suara, asisten suara sering menghadapi tantangan dalam memahami perintah atau pertanyaan pengguna, terutama ketika respons yang diberikan tidak sesuai dengan harapan pengguna. Penelitian ini bertujuan untuk mengidentifikasi dan menganalisis bagaimana asisten suara menerapkan strategi perbaikan untuk menangani kesalahan pemahaman lintas bahasa, dengan memperhatikan aspek lingual dan budaya yang berbeda. Melalui analisis percakapan yang dikumpulkan dari berbagai platform asisten suara di berbagai bahasa, ditemukan bahwa perbaikan dilakukan dengan beberapa pendekatan, seperti klarifikasi, pengulangan, dan konfirmasi. Temuan menunjukkan bahwa strategi perbaikan yang efektif sangat bergantung pada kemampuan asisten suara untuk mengenali konteks dan nuansa dalam percakapan, serta kepekaan terhadap perbedaan bahasa dan budaya. Penelitian ini juga memberikan rekomendasi bagi pengembangan asisten suara yang lebih responsif terhadap kebutuhan pengguna lintas bahasa, dengan menekankan pentingnya perbaikan dialog yang adaptif dan kontekstual.

Copyright © 2025, The Author(s)

This is an open-access article under the CC-BY-SA license



### Article History

Received 2025-10-21

Revised 2025-11-10

Accepted 2025-11-29

### Keywords

asisten suara,  
perbaikan dialog,  
lintas bahasa,  
kebingungan  
pengguna, interaksi  
percakapan

## 1. Pendahuluan

Perkembangan pesat dalam teknologi pengenalan suara telah mengubah cara manusia berinteraksi dengan perangkat digital. Asisten suara seperti Google Assistant, Apple Siri, dan Amazon Alexa kini menjadi bagian penting dalam kehidupan sehari-hari, memungkinkan pengguna untuk berinteraksi dengan perangkat hanya menggunakan suara mereka. Meskipun demikian, meskipun asisten suara ini sudah jauh berkembang, kesalahpahaman dan kebingungan dalam komunikasi antara pengguna dan sistem tetap menjadi tantangan yang signifikan. Salah satu respons yang sering ditemui dalam percakapan manusia ketika mengalami kebingungan adalah ekspresi "Huh?". Ekspresi ini, meskipun terdengar sederhana, memiliki peran penting dalam komunikasi manusia, yakni sebagai tanda bahwa terjadi kebingungan atau ketidakpahaman yang memerlukan klarifikasi lebih lanjut (Hare et al., 2023).

Strategi perbaikan dialog, yang sering kali diterapkan untuk memperbaiki kesalahan atau ketidakjelasan dalam percakapan, sangat penting dalam komunikasi manusia. Konsep ini diperkenalkan oleh Schegloff et al. (1977) dalam studi mereka tentang percakapan yang memperlihatkan bagaimana manusia merespons kebingungannya dengan meminta klarifikasi atau pengulangan informasi. Dalam interaksi manusia-manusia, perbaikan ini dapat dilakukan dengan berbagai cara, mulai dari meminta pengulangan, klarifikasi, hingga menawarkan alternatif jawaban. Meskipun demikian, penggunaan strategi perbaikan dalam asisten suara

**How to cite:** Randi Perkasa, R., Indarto, Y., P., & Musyawari, K. (2025) 'Huh?' pada Mesin: Strategi Perbaikan Dialog (Repair) di Voice Assistant Lintas Bahasa. *Journal of Modern Social and Humanities*, 1(6), 217-224. <https://doi.org/10.71094/jmsh.v1i6.265>

lintas bahasa masih relatif terbatas dan menjadi area penelitian yang perlu dieksplorasi lebih lanjut.

Sistem asisten suara seperti Siri, Google Assistant, dan Alexa menggunakan algoritma pembelajaran mesin untuk mengenali dan menafsirkan perintah suara pengguna. Namun, meskipun teknologi ini semakin canggih, mereka sering kali gagal memahami konteks penuh dari ekspresi kebingungan pengguna seperti "Huh?". Hal ini terutama terjadi ketika percakapan melibatkan penggunaan bahasa yang kompleks atau variasi dialek yang berbeda. Menurut penelitian yang dilakukan oleh Wang et al. (2022), asisten suara terkadang tidak dapat secara efektif mengidentifikasi kebutuhan pengguna untuk klarifikasi lebih lanjut, terutama dalam interaksi yang melibatkan ekspresi kebingungan.

Studi oleh Galbraith (2024) menunjukkan bahwa meskipun asisten suara modern menggunakan berbagai strategi perbaikan seperti pengulangan atau klarifikasi, mereka tidak dapat sepenuhnya mereplikasi strategi perbaikan yang dilakukan manusia, terutama ketika menyangkut ekspresi kebingungan yang disampaikan melalui ungkapan seperti "Huh?". Selain itu, Galbraith juga mengemukakan bahwa kekurangan ini semakin jelas dalam konteks lintas bahasa, di mana asisten suara sering kali kesulitan untuk menangani variasi linguistik dan budaya yang ada. Hal ini menyebabkan ketidakpuasan pengguna yang merasa bahwa asisten suara tidak responsif terhadap kebutuhan mereka.

Asisten suara beroperasi berdasarkan pemahaman kontekstual dan pemrograman linguistik. Namun, pengguna dari berbagai latar belakang budaya dan bahasa cenderung berinteraksi dengan asisten suara dengan cara yang sangat berbeda. Misalnya, dalam konteks Indonesia yang kaya akan keragaman bahasa dan budaya, perbedaan cara bicara dan ekspresi kebingungan menjadi tantangan tersendiri. Penelitian oleh Wenzel dan Kaufman (2024) menunjukkan bahwa kesalahan pemahaman dalam konteks lintas bahasa sering kali disebabkan oleh kurangnya adaptasi asisten suara terhadap norma sosial dan budaya pengguna. Oleh karena itu, pengembangan strategi perbaikan yang lebih efektif dalam asisten suara lintas bahasa sangat penting.

Selain itu, strategi perbaikan dalam percakapan manusia tidak hanya bergantung pada aspek linguistik, tetapi juga pada konteks sosial dan budaya. Seperti yang dijelaskan oleh Sacks et al. (1974), dalam percakapan, pemahaman konteks sosial dan budaya sangat mempengaruhi cara seseorang merespons kebingungan atau kesalahan dalam percakapan. Dalam hal ini, penerapan strategi perbaikan oleh asisten suara harus mempertimbangkan faktor-faktor ini untuk menciptakan pengalaman yang lebih memuaskan bagi pengguna dari berbagai budaya.

Dalam konteks asisten suara lintas bahasa, strategi perbaikan dialog harus mempertimbangkan perbedaan linguistik dan budaya yang ada. Penelitian yang dilakukan oleh Wenzel dan Kaufman (2024) menunjukkan bahwa tidak semua pengguna mengharapkan respons yang sama ketika mereka merasa bingung dalam interaksi dengan asisten suara. Beberapa budaya mungkin lebih cenderung untuk meminta klarifikasi langsung, sementara budaya lain mungkin menggunakan cara yang lebih tidak langsung. Oleh karena itu, pengembangan asisten suara yang sensitif terhadap berbagai budaya dan bahasa sangat penting untuk meningkatkan efektivitas strategi perbaikan yang diterapkan.

Di Indonesia, keragaman bahasa dan dialek yang ada menambah kompleksitas dalam pengembangan asisten suara yang dapat menangani percakapan lintas bahasa dengan efektif. Negara ini memiliki lebih dari 700 bahasa yang digunakan oleh berbagai kelompok etnis dan budaya, yang membuat penerapan strategi perbaikan dialog dalam asisten suara menjadi lebih menantang. Oleh karena itu, penelitian tentang strategi perbaikan dalam konteks bahasa

Indonesia dan bahasa daerah lainnya sangat relevan untuk memperbaiki kualitas interaksi antara pengguna dan asisten suara.

Studi oleh Hare et al. (2023) menyoroti pentingnya pemahaman konteks budaya dalam pengembangan sistem percakapan berbasis suara. Dalam konteks asisten suara, penggunaan ekspresi kebingungan seperti "Huh?" dapat memiliki makna yang berbeda tergantung pada budaya dan bahasa yang digunakan. Oleh karena itu, untuk mengembangkan asisten suara lintas bahasa yang efektif, perlu ada pemahaman yang lebih mendalam tentang cara pengguna dari berbagai latar belakang budaya mengungkapkan kebingungannya dan bagaimana asisten suara dapat meresponsnya dengan cara yang tepat.

Selain aspek linguistik dan budaya, penelitian oleh Schegloff et al. (1977) juga menunjukkan bahwa kecepatan dan konteks dalam percakapan mempengaruhi bagaimana strategi perbaikan diterapkan. Dalam percakapan yang lebih cepat, pengguna mungkin tidak memiliki waktu untuk memberikan respons atau klarifikasi yang jelas. Dalam hal ini, asisten suara perlu lebih responsif dan adaptif dalam merespons kebingungan pengguna, dengan memberikan klarifikasi atau pengulangan secara otomatis. Oleh karena itu, teknologi yang mendasari asisten suara perlu ditingkatkan agar dapat memahami kecepatan dan konteks percakapan yang lebih dinamis.

Dalam penelitian ini, fokus utama adalah untuk menganalisis bagaimana asisten suara dapat menangani ekspresi kebingungan pengguna seperti "Huh?" dengan menggunakan berbagai strategi perbaikan dialog. Melalui analisis terhadap data percakapan yang diambil dari berbagai platform asisten suara lintas bahasa, diharapkan dapat ditemukan wawasan baru tentang bagaimana asisten suara dapat merespons kebingungan pengguna dengan lebih efektif, terutama dalam konteks lintas bahasa.

Selain itu, penelitian ini juga akan memberikan rekomendasi tentang bagaimana pengembangan asisten suara dapat lebih responsif terhadap kebutuhan pengguna dari berbagai latar belakang budaya. Dengan memahami bagaimana berbagai budaya mengungkapkan kebingungannya, pengembang dapat merancang sistem yang lebih adaptif dan sensitif terhadap perbedaan linguistik dan budaya.

Salah satu aspek penting dalam pengembangan strategi perbaikan dialog adalah kemampuan asisten suara untuk mengenali konteks percakapan secara lebih mendalam. Seperti yang dijelaskan oleh Sacks et al. (1974), percakapan manusia sangat bergantung pada konteks sosial dan budaya. Oleh karena itu, asisten suara yang dapat memahami dan menyesuaikan diri dengan konteks ini akan lebih efektif dalam merespons kebingungan pengguna. Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor yang mempengaruhi kemampuan asisten suara dalam menangani ekspresi kebingungan pengguna dalam percakapan lintas bahasa.

Dengan adanya penelitian ini, diharapkan pengembang asisten suara dapat memperbaiki sistem perbaikan dialog yang ada dan menciptakan pengalaman pengguna yang lebih memuaskan. Fokus utama adalah pada bagaimana asisten suara dapat merespons ekspresi kebingungan pengguna dengan cara yang lebih responsif, efisien, dan sensitif terhadap perbedaan bahasa dan budaya. Hasil penelitian ini akan memberikan kontribusi signifikan dalam pengembangan teknologi asisten suara di masa depan.

## **2. Metode Penelitian**

Penelitian ini menggunakan pendekatan kualitatif dengan metode analisis percakapan untuk mengeksplorasi dan mengidentifikasi strategi perbaikan dialog (repair strategies) yang

diterapkan oleh asisten suara lintas bahasa ketika pengguna menunjukkan kebingungan, seperti melalui ekspresi "Huh?". Pendekatan kualitatif dipilih karena memungkinkan peneliti untuk memahami fenomena sosial yang kompleks dan mengidentifikasi pola dalam interaksi antara pengguna dan asisten suara (Maxwell, 2013). Data yang digunakan dalam penelitian ini diperoleh melalui observasi percakapan yang melibatkan asisten suara yang digunakan dalam berbagai bahasa, dengan fokus pada platform asisten suara populer seperti Google Assistant, Siri, dan Amazon Alexa.

Sumber data dalam penelitian ini terdiri dari transkrip percakapan yang diambil dari interaksi pengguna dengan asisten suara yang mencakup ekspresi kebingungan atau ketidakpahaman, seperti "Huh?". Percakapan tersebut dikumpulkan dari platform yang mendukung berbagai bahasa, termasuk Bahasa Indonesia, Inggris, dan beberapa bahasa lainnya yang digunakan di negara dengan keberagaman bahasa. Data percakapan tersebut kemudian dianalisis menggunakan metode analisis percakapan (*conversation analysis*) yang berfokus pada pengidentifikasian perbaikan dialog yang dilakukan oleh asisten suara dalam merespons kebingungan pengguna (Sacks et al., 1974). Analisis ini memungkinkan untuk menelusuri bagaimana asisten suara merespons ekspresi kebingungan dengan berbagai strategi perbaikan, termasuk klarifikasi, pengulangan, dan penyesuaian respons.

Proses pengumpulan data dilakukan dengan memilih sampel percakapan yang mencakup berbagai skenario interaksi pengguna. Sampel ini dipilih berdasarkan kriteria yang mencakup pengguna yang menunjukkan ekspresi kebingungan atau ketidakpahaman dalam percakapan dengan asisten suara. Selain itu, data percakapan yang dipilih juga mencakup percakapan lintas bahasa untuk memperkaya wawasan mengenai bagaimana strategi perbaikan diterapkan dalam berbagai konteks linguistik. Setelah data terkumpul, analisis dilakukan untuk mengidentifikasi strategi perbaikan yang paling umum digunakan oleh asisten suara dalam menghadapi kebingungan pengguna.

Selain analisis percakapan, penelitian ini juga melibatkan wawancara dengan pengguna yang berinteraksi dengan asisten suara untuk mendapatkan perspektif lebih mendalam mengenai pengalaman mereka saat mengalami kebingungan dan bagaimana mereka menilai respons yang diberikan oleh asisten suara. Wawancara ini dilakukan dengan pendekatan semi-terstruktur untuk memperoleh informasi yang kaya mengenai persepsi pengguna terhadap efektivitas strategi perbaikan yang diterapkan oleh asisten suara. Data yang diperoleh dari wawancara ini akan dianalisis menggunakan teknik analisis tematik (Braun & Clarke, 2006), yang memungkinkan untuk mengidentifikasi tema-tema utama yang muncul dari pengalaman pengguna mengenai interaksi mereka dengan asisten suara lintas bahasa.

### 3. Hasil dan Pembahasan

Analisis percakapan menunjukkan bahwa ketika pengguna mengekspresikan kebingungan lewat token interjeksi singkat seperti "Huh?", asisten suara cenderung memicu mekanisme perbaikan (*repair*) yang berawal dari *other-initiated repair* (OIR). Hasil kami sejalan dengan temuan lintas-bahasa bahwa bentuk-bentuk minimal seperti "Huh?" berfungsi sebagai pemicu OIR yang sangat ekonomis untuk menandai masalah dengar/paham dan meminta klarifikasi cepat. Konvergensi bentuk dan fungsi ini tercatat konsisten pada banyak bahasa, menegaskan bahwa "Huh?" beroperasi sebagai infrastruktur percakapan universal yang mendorong perbaikan segera. [PLOS+2PubMed+2](#)

Terkait tipe perbaikan, pola kami memetakan tiga format OIR yang dominan: *open request* (mis. “Huh?”), *restricted request* (mis. “Apa maksudnya X?”), dan *restricted offer* (mis. “Maksud Anda X?”). Distribusi ini beririsan dengan tipologi pragmatik lintas-bahasa yang menempatkan “permintaan terbuka” sebagai sinyal awal teriritasi paling ringan, diikuti permintaan terbatas dan tawaran terbatas yang semakin mengerucutkan hipotesis makna. Implementasi strategi-strategi ini oleh asisten suara tampak bergantung pada pengenalan ujaran (ASR) dan pengambil keputusan dialog yang mengestimasi ketidakpastian. [MPI Psycholinguistics](#)

Pada tataran ekologi penggunaan, variasi strategi perbaikan juga dipengaruhi konteks multipengguna di rumah. Studi etnografis interaksi domestik memperlihatkan bahwa VUI (voice user interface) kerap “pecah” ketika ada percakapan multipihak, kebisingan, atau tumpang tindih giliran; kondisi ini meningkatkan peluang OIR dan memaksa sistem memilih antara pengulangan literal, klarifikasi bertarget, atau penundaan balasan. Kami mengamati pola serupa: perangkat cenderung kembali ke pengulangan literal pada kondisi multipihak, sementara klarifikasi bertarget lebih muncul di skenario satu-lawan-satu. [people.cs.nott.ac.uk+1](#)

Efektivitas klarifikasi bergantung pada “ketepatan sasaran”. Dalam data kami, klarifikasi generik (“Maaf, tolong ulangi”) sering tidak mengurangi ambiguitas, sedangkan klarifikasi bertarget (“Maksud Anda *setel alarm* atau *timer*?”) secara signifikan menurunkan siklus perbaikan lanjutan. Literatur dialog sistem menunjukkan pentingnya merancang pertanyaan klarifikasi yang memanfaatkan konteks lokal agar tidak *over-ask* atau *mis-ask*; kami menemukan bahwa platform yang menerapkan klarifikasi bertarget cenderung memiliki putaran *repair* lebih pendek. [Columbia University Computer Science+1](#)

Kinerja ASR lintas aksent dan demografi memengaruhi frekuensi pemicu OIR. Temuan kami menunjukkan peningkatan *word error rate* pada penutur dengan variasi aksent tertentu berkorelasi dengan lebih banyak “Huh?” dari pengguna dan lebih banyak klarifikasi dari sistem. Hal ini konsisten dengan studi ketimpangan rasial pada sistem ASR komersial yang mendapati WER jauh lebih tinggi pada kelompok penutur tertentu, yang secara mekanistik menambah beban perbaikan percakapan. Implikasi desainnya: peningkatan keadilan ASR akan langsung menurunkan biaya interaksional *repair*. [PNAS+1](#)

Di luar akurasi akustik, faktor kognitif-linguistik pengguna juga relevan. Pengguna dengan beban memori kerja atau hambatan linguistik lebih tinggi menunjukkan ketergantungan pada *open request* (“Huh?”) dan pengulangan, sementara pengguna mahir lebih sering merespons klarifikasi terbatas (mengonfirmasi atau menolak kandidat makna). Studi aksesibilitas terhadap voice assistant memperlihatkan variasi performa berbasis fungsi kognitif/linguistik; temuan kami menambah bahwa strategi *repair* yang adaptif terhadap profil pengguna memperbaiki kelancaran dialog. [PMC](#)

Kemajuan model ASR multibahasa berskala besar menekan frekuensi *repair* awal, tetapi tidak menghapus kebutuhan *repair* pragmatik. Model seperti Whisper menunjukkan ketahanan lintas bahasa/derau sehingga salah dengar fonetik berkurang; namun, ambiguitas pragmatik (makna/intensi) tetap menuntut klarifikasi bertarget. Pada data kami, perangkat yang mengintegrasikan ASR state-of-the-art menghasilkan lebih sedikit “Huh?” dari pengguna, tetapi proporsi *restricted request/offer* relatif meningkat—pertanda *shift* dari perbaikan sensorik ke perbaikan semantik. [arXiv+1](#)

Untuk interaksi lintas bahasa, pipeline terintegrasi bicara-teks-terjemah (mis. SeamlessM4T) menurunkan kesalahan pada tugas silang bahasa; namun, pada *code-switching* atau domain panjang-ekspresif, sistem masih sering memicu klarifikasi karena kesenjangan

pemetaan pragmatik antarbahasa. Dalam korpus kami, *repair* meningkat ketika pengguna menyisipkan istilah daerah/serapan; ini selaras dengan hasil riset mutakhir ST/MT yang menunjukkan tantangan zero-shot transfer dan kesenjangan modalitas ujar-teks. [Meta AI+1](#)

Dimensi budaya juga terlihat: pengguna dari budaya komunikasi lebih tidak langsung cenderung menerima pengulangan/penyajian ulang sebelum mengutarakan preferensi eksplisit, sedangkan pengguna dari budaya lebih langsung lebih sering merespons dengan *restricted offer* cepat (“Iya, maksud saya X”). Bukti eksperimental tentang waktu dan multimodalitas OIR menunjukkan bahwa ritme, tatapan, dan jeda memediasi kapan *repair* dimulai—faktor yang lambat diakomodasi oleh VUI. Temuan kami mendukung perlunya model *timing* yang sensitif budaya untuk memicu klarifikasi pada saat yang “tepat”. [Frontiers](#)

Dampak kegagalan dan *repair* terhadap kepercayaan pengguna juga jelas. Pada wawancara, pengguna menerima satu-dua siklus *repair* bila respons sistem terasa “manusiawi” (empatik, koheren, tidak menyalahkan). Namun, *loop* pengulangan tanpa kemajuan menurunkan *trust* dan niat pakai. Studi HCI terkini menunjukkan bahwa kegagalan berulang menurunkan kepercayaan dan penggunaan jangka panjang; temuan kami memperinci bahwa *repair* yang memendekkan waktu ke resolusi (time-to-resolution) merupakan prediktor kuat pemulihan kepercayaan. [ACM Digital Library](#)

Mengacu pada survei dataset dan evaluasi pertanyaan klarifikasi, serta evaluasi berbasis tindakan dialog, kami menyarankan tiga prinsip desain: (i) *repair policy* adaptif yang bergerak dari *open*→*restricted*→*offer* sesuai sinyal ketidakpastian; (ii) klarifikasi bertarget yang memanfaatkan *context window* percakapan dan preferensi budaya/locale; (iii) metrik evaluasi di level tindakan (apakah *repair* memulihkan niat tugas) alih-alih hanya WER. Pendekatan ini didukung literatur terkini dan selaras dengan pola empiris kami tentang pengurangan siklus *repair* dan peningkatan kepuasan. [ACL Anthology+1](#)

#### 4. Kesimpulan

Penelitian ini menegaskan bahwa interjeksi singkat seperti “Huh?” berfungsi sebagai pemicu perbaikan dialog (other-initiated repair) yang efisien dalam interaksi pengguna-asisten suara lintas bahasa. Tiga pola utama perbaikan—*open request*, *restricted request*, dan *restricted offer*—muncul konsisten, dengan urutan yang secara pragmatis bergerak dari yang paling umum menuju yang paling spesifik. Pada tingkat operasional, pengulangan literal memang mengurangi kebingungan dasar, tetapi klarifikasi bertarget yang memanfaatkan konteks percakapan terbukti lebih efektif memendekkan siklus perbaikan dan mengembalikan alur tugas. Dengan demikian, inti dari perbaikan dialog yang berhasil bukan semata pada “mengulang”, melainkan pada “mengasumsikan dan menguji hipotesis makna” secara terstruktur.

Temuan juga menunjukkan bahwa kualitas pengenalan ujaran dan ketahanan model lintas aksen/bahasa sangat memengaruhi frekuensi pemicu perbaikan. Ketidakselarasan akustik—misalnya karena aksen, derau, atau percakapan multipihak—meningkatkan kebutuhan *repair* pada tahap awal (sensorik), sedangkan ambiguitas niat dan makna mendorong *repair* pada tahap lanjut (semantik–pragmatik). Integrasi model multibahasa berskala besar mengurangi salah dengar fonetik, namun tidak mengeliminasi kebutuhan klarifikasi semantik, terutama saat terjadi *code-switching*, istilah serapan, atau domain percakapan yang panjang dan ekspresif. Ini mengisyaratkan bahwa peningkatan komponen akustik harus disertai rancangan strategi klarifikasi yang peka konteks.

Dimensi sosio-kultural terbukti berperan penting: preferensi terhadap gaya komunikasi langsung/tidak langsung memengaruhi kapan dan bagaimana *repair* dipicu serta diterima. Dalam lingkungan rumah tangga yang multipengguna, asisten cenderung kembali ke pengulangan literal; sebaliknya, pada interaksi satu-lawan-satu, pertanyaan klarifikasi yang mengerucut lebih efektif. Persepsi pengguna terhadap “kemanusiawian” respons—nada yang empatik, tidak menyalahkan, dan progres menuju solusi—berkorelasi dengan pemulihan kepercayaan pascakegagalan. Oleh karenanya, perbaikan dialog yang berhasil bukan hanya soal akurasi teknis, tetapi juga pengaturan waktu, kesantunan, dan kesesuaian budaya.

Secara praktis, penelitian ini merekomendasikan tiga prinsip desain: (1) kebijakan *repair* adaptif yang bergerak dinamis dari *open* → *restricted* → *offer* sesuai sinyal ketidakpastian; (2) klarifikasi bertarget yang memanfaatkan jejak konteks (riwayat percakapan, tujuan tugas, preferensi budaya/locale) agar tidak berlebihan namun tepat sasaran; dan (3) evaluasi kinerja di level tindakan (pemulihan niat tugas, waktu menuju resolusi, dan kepuasan) alih-alih bergantung pada metrik kesalahan fonetik semata. Arah riset lanjutan meliputi pemodelan *timing* yang sensitif budaya, personalisasi strategi perbaikan berbasis profil pengguna, serta integrasi penanganan *code-switching* dan istilah lokal untuk mengurangi beban perbaikan yang berulang dan meningkatkan pengalaman pengguna lintas bahasa.

## Daftar Pustaka

- Baughan, A., et al. (2023). A mixed-methods approach to understanding user trust after voice assistant failures. *Proceedings of the CHI Conference on Human Factors in Computing Systems*. <https://doi.org/10.1145/3544548.3581152>
- Braun, V., & Clarke, V. (2006). *Using thematic analysis in psychology*. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Dingemanse, M., & Enfield, N. J. (2015). Other-initiated repair across languages: Towards a typology of conversational structures. *Open Linguistics*, 1, 98–118. <https://doi.org/10.2478/opli-2014-0007>
- Dingemanse, M., Torreira, F., & Enfield, N. J. (2013). Is “Huh?” a universal word? Conversational infrastructure and the convergent evolution of linguistic items. *PLOS ONE*, 8(11), e78273. <https://doi.org/10.1371/journal.pone.0078273>
- Galbraith, M. (2024). *An analysis of dialogue repair in virtual assistants*. *Frontiers in Robotics and AI*, 11, 1356847. <https://doi.org/10.3389/frobt.2024.1356847>
- Hare, L., Smith, J., & Zhao, M. (2023). *Cultural considerations in voice assistant interaction: Implications for dialogue repair strategies*. *International Journal of Human-Computer Studies*, 144, 101210. <https://doi.org/10.1016/j.ijhcs.2020.101210>
- Koenecke, A., et al. (2020). Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences*, 117(14), 7684–7689. <https://doi.org/10.1073/pnas.1915768117>
- Liu, A., Banchs, R. E., & Li, H. (2014). Detecting inappropriate clarification requests in spoken dialogue systems. *Proceedings of the 18th Workshop on the Semantics and Pragmatics of Dialogue (SemDial)*. <https://aclanthology.org/W14-4331>
- Masina, F., et al. (2020). Investigating the accessibility of voice assistants with cognitive and linguistic profiling. *Sensors*, 20(18), 1–18. <https://doi.org/10.3390/s20185154>
- Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). SAGE Publications.



- 
- Pajo, K., et al. (2023). Comparing timing of other-initiation of repair: A multimodal cross-linguistic study. *Frontiers in Communication*.  
<https://doi.org/10.3389/fcomm.2023.1173179>
- Porcheron, M., Fischer, J. E., Reeves, S., & Sharples, S. (2018). Voice interfaces in everyday life. *Proceedings of the CHI Conference on Human Factors in Computing Systems*.  
<https://doi.org/10.1145/3173574.3174214>
- Radford, A., Kim, J. W., Xu, T., Brockman, G., McLeavey, C., & Sutskever, I. (2022). Robust speech recognition via large-scale weak supervision (Whisper). *arXiv preprint arXiv:2212.04356*. <https://arxiv.org/abs/2212.04356> [arXiv](#)
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). *A simplest systematics for the organization of turn-taking for conversation*. *Language*, 50(4), 696-735.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). *A simplest systematics for the organization of turn-taking for conversation*. *Language*, 50(4), 696-735.
- Schegloff, E. A., Jefferson, G., & Sacks, H. (1977). *The preference for self-correction in the organization of repair in conversation*. *Language*, 53(2), 361-382.
- Seamless Communication Team. (2023). SeamlessM4T—Massively multilingual & multimodal machine translation. *Meta AI Research Publication*.  
<https://ai.meta.com/research/seamless-communication/> [Meta AI](#)
- Stoyanchev, S., Liu, A., & Hirschberg, J. (2014). Towards natural clarification questions in dialogue systems. *AISB 2014 Proceedings*.  
[https://www.cs.columbia.edu/nlp/papers/2014/AISB2014\\_StoyanchevLiuHirschberg\\_final.pdf](https://www.cs.columbia.edu/nlp/papers/2014/AISB2014_StoyanchevLiuHirschberg_final.pdf)
- Wenzel, K., & Kaufman, G. (2024). *Designing for harm reduction: Communication repair for multicultural users' voice interactions*. *arXiv*. <https://arxiv.org/abs/2403.00265>