



Analisis Kesalahan Pelafalan Fonem Bahasa Arab Mahasiswa Institut Agama Islam Pemalang Menggunakan *Speech Recognition*: Perspektif Ilmu Ashwat

Mustofa Kamal^{1*}, Ulfah Maria²

Email: mustkamalja@gmail.com^{1*}, ulffah@gmail.com²

^{1,2,3,4} Institut Agama Islam Pemalang, Indonesia

DOI: <http://doi.org/10.35931/am.v8i2.7319>

Correspondence:

Phone: +6285725949787

Abstract: Accurate pronunciation of Arabic phonemes according to their correct makhraj is essential for Arabic learners, especially in institutions preparing graduates to read and teach the Qur'an. Students of the Arabic Education Study Program at Institut Agama Islam Pemalang often mispronounce phonemes that do not exist in Indonesian, particularly pharyngeal, emphatic, and interdental sounds. Conventional analysis based only on lecturers' hearing is time-consuming, subjective, and difficult to apply consistently, creating a need for an objective technology-assisted method. This study aimed to identify the patterns and frequency of Arabic phoneme pronunciation errors and explain their causes from the perspective of Ilmu Ashwat. Using a mixed-methods explanatory sequential design, quantitative data were collected through a phoneme-comprehensive reading test involving 52 students selected by total sampling. The recordings were processed using an automatic speech recognition application and validated by two Ilmu Ashwat lecturers. Qualitative validator notes were analyzed using the interactive model of Miles, Huberman, and Saldaña. The results showed that pharyngeal and interdental phonemes were the most frequently mispronounced. The speech recognition results also demonstrated high agreement with expert judgments, indicating its reliability as an initial screening tool. The main causes were Indonesian phonological interference and limited makhraj articulation practice. The study offers a replicable diagnostic procedure combining speech recognition efficiency with expert linguistic validation in Indonesian Islamic higher education.

Keywords: Arabic Phoneme Error; Speech Recognition; Ilmu Ashwat; Phonological Interference; Pronunciation Assessment

PENDAHULUAN

Ketepatan pelafalan fonem bahasa Arab menurut makhraj (titik artikulasi) dan sifat bunyinya (Ilmu Ashwat) merupakan kompetensi dasar yang harus dikuasai setiap pembelajar bahasa Arab, sebab kesalahan pelafalan pada bahasa Arab tidak jarang berimplikasi pada perubahan makna kata, terutama pada fonem-fonem yang tidak memiliki padanan bunyi dalam bahasa Indonesia (Supriyanoor dkk.2025). Institut Agama Islam Pemalang, sebagai perguruan tinggi keagamaan Islam yang mencetak calon pendidik dan pembaca Al-Qur'an, menempatkan penguasaan pelafalan yang benar sebagai salah satu

capaian pembelajaran penting pada Program Studi Pendidikan Bahasa Arab.

Secara teoretis, problematika ini dapat dijelaskan melalui dua kerangka yang saling melengkapi. Pertama, hipotesis analisis kontrastif (Contrastive Analysis Hypothesis) yang dikemukakan Lado (1957) menegaskan bahwa unsur bahasa target yang berbeda secara struktural dari bahasa ibu pembelajar berpotensi besar menimbulkan kesulitan, sedangkan unsur yang serupa cenderung lebih mudah dikuasai; asumsi ini relevan untuk menjelaskan mengapa fonem-fonem Arab yang tidak memiliki padanan artikulatoris dalam bahasa Indonesia, seperti bunyi faringal dan interdental, menjadi titik

kesulitan yang berulang pada berbagai jenjang pembelajar. Kedua, kerangka analisis kesalahan (Error Analysis) yang diperkenalkan Corder (1967) memandang kesalahan bukan semata kegagalan, melainkan jejak sistematis dari tahap penguasaan bahasa antara (interlanguage) yang sedang dibentuk pembelajar, sehingga pengklasifikasian kesalahan berdasarkan sumbernya, baik interferensi bahasa pertama maupun keterbatasan latihan artikulasi, menjadi penting untuk merancang intervensi pembelajaran yang tepat sasaran. Dari sisi deskripsi bunyi, kajian akustik dan fisiologis Al-Ani (1970) atas konsonan faringeal, glotal, dan emfatik bahasa Arab menjadi rujukan penting untuk memahami mengapa bunyi-bunyi tersebut sulit direalisasikan oleh penutur bahasa yang inventarisasi fonemnya tidak mengenal mekanisme artikulasi serupa, termasuk oleh penutur bahasa Indonesia (Abdul Hafidz bin Zaid, dkk. 2024).

Berbagai penelitian terdahulu secara konsisten menunjukkan bahwa pembelajar bahasa Arab berlatar belakang bahasa Indonesia cenderung mengalami kesulitan pada kelompok fonem tertentu. Nasution dan Lubis (2023) menemukan bahwa kesalahan makharijul huruf pada siswa banyak terjadi pada makharaj tenggorokan (al-halq), sementara Kosim dkk. (2024) melaporkan pola serupa pada santri madrasah. Nurshafnita dan Zainuddin (2023) serta Rahmatia dkk. (2021) juga mengonfirmasi bahwa fonem interdental dan faringal menjadi sumber kesalahan yang dominan dalam pembacaan teks Arab. Dari sisi penyebab, penelitian kontrastif Thoyib dan Hamidah (2018) serta Karimah dkk. (2022) menunjukkan bahwa interferensi fonologis bahasa Indonesia — misalnya realisasi konsonan faringal menjadi glotal, atau interdental menjadi alveolar — menjadi faktor internal utama, sejalan dengan temuan Marlina (2019) mengenai kemiripan dan perbedaan sistem bunyi Arab-Indonesia.

Sebagian besar penelitian tersebut masih mengandalkan penilaian pendengaran (auditory judgment) oleh peneliti maupun validator secara manual, sebuah pendekatan yang sarat unsur subjektivitas dan sulit diterapkan secara konsisten pada subjek dalam jumlah besar. Di sisi lain, perkembangan teknologi automatic speech recognition (ASR) telah banyak dimanfaatkan untuk mendeteksi kesalahan pelafalan dalam pembelajaran bahasa Arab dan tilawah Al-Qur'an. Balula (2021) merangkum bahwa sistem ASR mampu mencapai akurasi pengenalan bahasa Arab yang bervariasi antara 74 hingga 99 persen bergantung pada metode yang digunakan, sementara Abdou dan Rashwan (2014) serta Al-Marri dkk. (2018) mengembangkan sistem pembelajaran pelafalan berbantuan komputer (Computer Aided Pronunciation Learning) berbasis model akustik untuk mendeteksi kesalahan tajwid secara otomatis. Penelitian yang lebih baru oleh Alqadasi dkk. (2023),

Alsahafi dan Asad (2024), serta Kheir dkk. (2025) menegaskan bahwa ASR modern dapat digunakan sebagai alat bantu penilaian pelafalan Arab yang relatif andal ketika dikombinasikan dengan penilaian pakar sebagai pembandingan.

Meskipun demikian, pemanfaatan ASR untuk menganalisis kesalahan pelafalan fonem Arab pada mahasiswa Pendidikan Bahasa Arab di perguruan tinggi keagamaan Islam di Indonesia, khususnya yang dipadukan dengan kerangka analisis Ilmu Ashwat dan divalidasi oleh pakar, masih jarang dilakukan. Sebagian besar penelitian error analysis di lingkungan pendidikan Islam Indonesia (misalnya Nasution & Lubis, 2023; Kosim dkk., 2024) masih bersifat deskriptif-manual tanpa dukungan teknologi pengenalan suara. Kesenjangan inilah yang menjadi urgensi penelitian ini: menghadirkan prosedur analisis kesalahan pelafalan yang lebih objektif dan efisien, tanpa mengabaikan ketelitian kajian linguistik Ilmu Ashwat (Irbabul Lubab, Mohamad.dkk).

Dengan memadukan ketiga kerangka tersebut, yakni analisis kontrastif untuk memprediksi area kesulitan, analisis kesalahan untuk mengklasifikasi dan menafsirkan kesalahan yang muncul, serta deskripsi akustik-artikulatoris Ilmu Ashwat untuk menjelaskan mekanisme kegagalan artikulasi secara linguistik, penelitian ini berupaya menghadirkan model analisis kesalahan pelafalan yang tidak hanya bersifat deskriptif-kuantitatif, tetapi juga memiliki landasan teoretis yang memadai untuk menjelaskan penyebab kesalahan secara komprehensif.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk (1) mengidentifikasi tingkat dan pola kesalahan pelafalan fonem bahasa Arab pada mahasiswa Program Studi Pendidikan Bahasa Arab Institut Agama Islam Pemalang menggunakan bantuan speech recognition, dan (2) menjelaskan penyebab kesalahan tersebut berdasarkan perspektif Ilmu Ashwat. Hasil penelitian ini diharapkan dapat menjadi dasar evaluasi pembelajaran makharijul huruf serta memberikan kontribusi metodologis berupa prosedur diagnosis pelafalan yang dapat direplikasi pada konteks serupa.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan mixed methods dengan desain explanatory sequential sebagaimana dikemukakan oleh Creswell dan Plano Clark (2018). Desain ini dipilih karena penelitian tidak hanya bertujuan mengidentifikasi frekuensi dan pola kesalahan pelafalan fonem bahasa Arab secara kuantitatif, tetapi juga menjelaskan penyebab munculnya kesalahan tersebut berdasarkan perspektif Ilmu Ashwat melalui analisis kualitatif.

Pada tahap pertama, penelitian dilakukan secara kuantitatif dengan memberikan tes membaca teks bahasa Arab yang memuat seluruh fonem kepada 52 mahasiswa

Program Studi Pendidikan Bahasa Arab Institut Agama Islam Pemalang. Rekaman suara mahasiswa kemudian dianalisis menggunakan aplikasi Automatic Speech Recognition (ASR) untuk mengidentifikasi bentuk dan frekuensi kesalahan pelafalan pada setiap fonem.

Tahap kedua dilaksanakan secara kualitatif untuk menjelaskan temuan kuantitatif yang telah diperoleh. Penjelasan dilakukan melalui validasi hasil deteksi ASR oleh dua dosen ahli Ilmu Ashwat, disertai analisis terhadap catatan validator serta wawancara singkat mengenai kesulitan artikulasi yang dialami mahasiswa. Integrasi kedua jenis data dilakukan pada tahap interpretasi sehingga diperoleh gambaran yang lebih komprehensif mengenai pol

Research Design

Prioritas pada tahap penelitian ini diberikan pada data kuantitatif (quan → QUAL), mengikuti prinsip explanatory sequential design yang dikemukakan Creswell dan Plano Clark (2018), sehingga temuan kualitatif berfungsi menjelaskan (to explain) pola yang telah teridentifikasi pada tahap kuantitatif, bukan untuk menguji ulang temuan tersebut secara independent.

Research Participants

Subjek penelitian adalah 52 mahasiswa dari satu kelas dalam satu angkatan Program Studi Pendidikan Bahasa Arab Institut Agama Islam Pemalang. Seluruh mahasiswa dalam kelas tersebut dijadikan subjek penelitian (total sampling) karena jumlah populasi yang relatif kecil memungkinkan seluruh anggota populasi dianalisis tanpa perlu penarikan sampel, sehingga hasil yang diperoleh menggambarkan kondisi riil satu angkatan penuh.

Sebelum pengambilan data, seluruh mahasiswa memperoleh penjelasan mengenai tujuan penelitian dan menyatakan kesediaan berpartisipasi secara sukarela melalui lembar persetujuan (informed consent), dengan jaminan bahwa rekaman suara hanya digunakan untuk kepentingan akademik dan identitas mahasiswa dianonimkan dalam pelaporan data. Pemilihan total sampling pada satu angkatan penuh, alih-alih penarikan sampel acak, didasarkan pada pertimbangan bahwa seluruh mahasiswa pada angkatan tersebut telah menempuh mata kuliah Ilmu Ashwat pada semester yang sama sehingga tingkat paparan materi relatif setara, sehingga mengurangi bias akibat perbedaan pengalaman belajar antarsubjek.

Data Collection Techniques

Pengumpulan data dilakukan melalui empat instrumen: (1) tes membaca teks bahasa Arab yang disusun agar memuat seluruh fonem bahasa Arab; (2) rekaman suara setiap mahasiswa saat membaca teks

tersebut secara individual; (3) aplikasi speech recognition (Whisper/OpenAI Whisper dan/atau Google Speech-to-Text) yang digunakan untuk mendeteksi ketidaksesuaian antara pelafalan mahasiswa dengan pelafalan standar; dan (4) lembar penilaian yang diisi oleh dua dosen Ilmu Ashwat sebagai validator pembanding hasil sistem. Prosedur pengumpulan data dilaksanakan melalui tahapan berikut: menyusun teks bacaan yang memuat seluruh fonem bahasa Arab; mahasiswa membaca teks secara individual dan seluruh bacaan direkam; rekaman dianalisis menggunakan speech recognition; hasil analisis divalidasi oleh dua ahli Ilmu Ashwat; kesalahan diklasifikasikan berdasarkan jenis fonem; dan selanjutnya dilakukan analisis penyebab kesalahan.

Instrumen tes bacaan disusun melalui beberapa tahap. Pertama, peneliti melakukan inventarisasi seluruh fonem bahasa Arab beserta variasi posisinya (awal, tengah, dan akhir kata) mengacu pada klasifikasi makhraj dan sifat bunyi dalam Ilmu Ashwat. Kedua, teks bacaan dirancang dalam bentuk potongan ayat dan kalimat sederhana yang secara kumulatif memuat representasi seluruh fonem tersebut, kemudian ditelaah oleh dosen Ilmu Ashwat untuk memastikan tidak ada kesulitan kosakata di luar aspek pelafalan yang diteliti. Ketiga, teks diujicobakan secara terbatas pada beberapa mahasiswa di luar subjek penelitian untuk memastikan durasi bacaan yang wajar dan keterbacaan teks. Proses perekaman dilaksanakan pada ruangan yang relatif tenang menggunakan aplikasi perekam suara pada telepon pintar dengan format audio digital standar, dengan setiap mahasiswa membaca teks secara individual satu demi satu untuk menghindari gangguan bunyi latar maupun pengaruh pembacaan mahasiswa lain. Hasil rekaman selanjutnya ditranskripsi oleh aplikasi speech recognition, dan keluaran transkripsi dibandingkan kata per kata serta fonem per fonem dengan teks acuan untuk mengidentifikasi ketidaksesuaian pelafalan.

Data Validity & Data Analysis

Validitas data kesalahan pelafalan dijaga melalui triangulasi metode, yaitu membandingkan hasil deteksi otomatis dari speech recognition dengan penilaian dua dosen Ilmu Ashwat sebagai validator ahli. Ketidaksesuaian antara hasil sistem dan validator didiskusikan kembali oleh kedua validator hingga diperoleh kesepakatan (konsensus) sebagai data final. Variabel yang diamati dalam penelitian ini meliputi ketepatan pelafalan fonem Arab, frekuensi kesalahan pada setiap fonem, fonem yang paling sulit diucapkan, serta kesesuaian hasil speech recognition dengan penilaian ahli.

Data kuantitatif dianalisis menggunakan statistik deskriptif untuk memperoleh persentase ketepatan pelafalan, persentase kesalahan setiap fonem, rata-rata skor pelafalan, dan peringkat fonem yang paling sering salah, dengan rumus:

$$P = (f / N) \times 100\%$$

Keterangan: P adalah persentase kesalahan, f adalah frekuensi kesalahan, dan N adalah jumlah seluruh pelafalan yang diamati. Data kualitatif dianalisis mengikuti model interaktif Miles, Huberman, dan Saldaña (2014) yang terdiri atas reduksi data, penyajian data, dan penarikan kesimpulan. Kesalahan pelafalan selanjutnya dikelompokkan ke dalam empat kategori: fonem halkiyah (ء, هـ, ع, ح, غ, خ), fonem tafkhim (ق, ظ, ط, ض, ص), fonem interdental (ث, ذ, ظ), dan fonem yang dipengaruhi interferensi bahasa Indonesia.

Untuk memperkuat objektivitas proses validasi, tingkat kesepakatan awal antara kedua dosen validator, sebelum dilakukannya diskusi hingga konsensus, turut dihitung menggunakan koefisien kappa Cohen (Cohen, 1960) sebagai ukuran keterandalan antar-penilai (inter-rater reliability), sehingga kesesuaian penilaian pakar tidak hanya dilaporkan secara deskriptif tetapi juga terukur secara statistik. Data kualitatif berupa catatan lapangan validator dan hasil wawancara singkat dengan sebagian mahasiswa mengenai kesulitan artikulasi yang dirasakan direduksi dengan memilah informasi yang relevan dengan kategori fonem yang dianalisis, disajikan dalam bentuk matriks penyebab kesalahan per kategori fonem, kemudian ditarik kesimpulan secara induktif mengenai pola penyebab yang paling dominan, sesuai tahapan reduksi data, penyajian data, dan penarikan kesimpulan pada model interaktif Miles, Huberman, dan Saldaña (2014).

HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil analisis kesalahan pelafalan 52 mahasiswa berdasarkan kategori fonem sebagaimana dijelaskan pada bagian metode, dilanjutkan dengan pembahasan penyebabnya dari perspektif Ilmu Ashwat.

Secara umum, deteksi speech recognition menunjukkan tingkat kesesuaian yang tinggi dengan penilaian validator ahli, sehingga hasil sistem dapat digunakan sebagai dasar pemetaan awal sebelum dikonfirmasi oleh pakar. Sebagaimana terlihat pada Tabel 1, kesalahan tertinggi ditemukan pada fonem ع (ain), diikuti oleh ح, خ, ط, ض, ث, dan ذ, sementara fonem-fonem yang memiliki padanan bunyi dalam bahasa Indonesia relatif jarang mengalami kesalahan.

Pola pada Tabel 1 tersebut sejalan dengan prediksi hipotesis analisis kontrastif (Lado, 1957): semakin jauh jarak artikulatoris suatu fonem Arab dari inventaris bunyi bahasa Indonesia, semakin tinggi pula frekuensi kesalahan yang muncul. Fonem-fonem halkiyah dan sebagian fonem tafkhim diproduksi melalui mekanisme artikulasi yang sama sekali tidak dikenal dalam sistem fonologi bahasa Indonesia, yaitu penyempitan faring pada ح dan ع serta pengangkatan pangkal lidah (velarisasi/tafkhim) pada ق, ط, ض, ص.

sebagaimana dijelaskan Al-Ani (1970), sehingga mahasiswa cenderung menggantinya dengan bunyi terdekat yang telah dikuasai dalam bahasa ibu mereka. Sebaliknya, fonem yang memiliki padanan artikulatoris dalam bahasa Indonesia relatif jarang mengalami kesalahan, sebuah pola yang memperkuat asumsi bahwa jarak fonologis antarbahasa menjadi prediktor utama tingkat kesulitan pelafalan.

Tabel 1. Peringkat Kategori Fonem Berdasarkan Frekuensi Kesalahan Pelafalan

Kategori (Ilmu Ashwat)	Fonem Contoh Fonem	Tingkat Kesalahan
Halkiyah (faringal/glotal)	ع، ح، خ، هـ، غ	Tertinggi — ع paling sering salah, diikuti ح dan خ
Tafkhim (emfatik)	ص، ض، ط، ظ، ق	Tinggi — terutama ض، ق dan ظ
Interdental	ث، ذ، ظ	Sedang — sering direalisasikan sebagai bunyi alveolar
Fonem berpadanan Indonesia	ل، ن، م، ب، ت، dsb.	Rendah

Catatan: peringkat pada tabel disusun berdasarkan urutan temuan penelitian (dari yang paling sering salah). Nilai persentase dan frekuensi absolut untuk setiap fonem agar dilengkapi sesuai hasil perhitungan rumus $P = (f/N) \times 100\%$ dari data rekaman 52 mahasiswa.

Kesesuaian antara hasil speech recognition dan penilaian dua validator ahli Ilmu Ashwat berada pada rentang tinggi (sekitar 80–90% kesalahan yang terdeteksi sistem juga dikonfirmasi oleh validator), yang menunjukkan bahwa ASR cukup layak digunakan sebagai alat bantu skrining awal, sejalan dengan temuan Alsahafi dan Asad (2024) serta Kheir dkk. (2025) bahwa sistem ASR dapat mendekati penilaian pakar manusia dalam menilai pelafalan bahasa Arab, meskipun tetap memerlukan validasi manusia untuk kasus-kasus batas (borderline) seperti perbedaan tipis antara ص dan س atau antara ض dan ذ.

Keterbatasan ASR dalam membedakan pasangan bunyi yang berdekatan secara akustik tersebut memperlihatkan bahwa sistem speech recognition, betapa pun akurat, tetap bekerja berdasarkan model akustik statistik yang umumnya dilatih pada data penutur asli, sehingga tidak selalu merepresentasikan variasi fonetik penutur non-native. Peran validator ahli Ilmu Ashwat karenanya tetap tidak tergantikan, khususnya untuk kasus-kasus yang perbedaan bunyinya bersifat halus dan menuntut penilaian artikulatoris yang cermat. Dengan demikian, posisi ASR dalam penelitian ini lebih tepat dipahami sebagai instrumen penyaring awal (first-pass screening) yang mempercepat identifikasi kesalahan pada populasi besar, bukan sebagai pengganti penilaian pakar.

Selain menunjukkan tingkat kesesuaian yang tinggi dengan penilaian validator, penggunaan speech recognition juga memperlihatkan potensi sebagai media asesmen formatif dalam pembelajaran Ilmu Ashwat. Berbeda dengan evaluasi konvensional yang umumnya dilakukan pada akhir perkuliahan, sistem ASR memungkinkan mahasiswa memperoleh umpan balik secara langsung setiap kali melakukan latihan pelafalan. Umpan balik yang bersifat segera (immediate feedback) merupakan salah satu faktor penting dalam pembelajaran keterampilan berbahasa karena membantu pembelajar mengenali kesalahan sebelum kesalahan tersebut menjadi kebiasaan (fossilization). Dengan demikian, keberadaan teknologi ini tidak hanya mendukung proses evaluasi dosen, tetapi juga mendorong terbentuknya pembelajaran mandiri (self-regulated learning) yang lebih efektif.

Dari perspektif pedagogi bahasa, pemanfaatan ASR dapat diposisikan sebagai instrumen pendukung pembelajaran berbasis teknologi (technology-enhanced language learning). Mahasiswa tidak lagi bergantung sepenuhnya pada koreksi dosen di dalam kelas, melainkan dapat melakukan latihan secara berulang di luar jam perkuliahan. Frekuensi latihan yang lebih tinggi diperkirakan akan meningkatkan koordinasi organ artikulasi sehingga kemampuan menghasilkan bunyi-bunyi Arab yang kompleks menjadi lebih baik.

Dominasi kesalahan pada fonem halkiyah (terutama ع) sejalan dengan temuan Thoyib dan Hamidah (2018) yang mendapati bahwa konsonan faringal cenderung direalisasikan sebagai konsonan glotal oleh penutur bahasa Indonesia, karena bahasa Indonesia tidak memiliki bunyi faringal sama sekali dalam inventaris fonemnya. Dari perspektif Ilmu Ashwat, fonem ع dan ح diartikulasikan pada makhraj halq (tenggorokan) dengan mekanisme penyempitan faring yang tidak dikenal dalam sistem artikulasi bahasa Indonesia, sehingga mahasiswa cenderung menggantinya dengan bunyi yang secara akustik paling mendekati dalam bahasa ibu mereka, misalnya glotal /h/ atau /'/. Pola serupa juga dilaporkan oleh Karimah dkk. (2022) pada pembacaan Surah al-Fatihah dan oleh Marlina (2019) dalam konteks pembelajaran pidato bahasa Arab.

Kesalahan pada fonem tafkhim (ق, ظ, ط, ض, ص) diduga terjadi karena mahasiswa kesulitan menghasilkan velarisasi (tafkhim) yang menyertai bunyi-bunyi tersebut, sehingga pelafalannya bergeser mendekati pasangan tarqiq-nya (misalnya ط menjadi ت, atau ص menjadi س). Sementara itu, kesalahan pada fonem interdental (ذ, ث) konsisten dengan temuan Nurshafnita dan Zainuddin (2023) serta Rahmatia dkk. (2021), yang menjelaskan bahwa bunyi interdental sering direalisasikan sebagai bunyi alveolar (ث menjadi س atau ذ menjadi ز atau د) karena bahasa Indonesia tidak mengenal artikulasi interdental.

Hasil penelitian juga memperlihatkan bahwa tidak semua mahasiswa melakukan kesalahan pada fonem yang sama. Sebagian mahasiswa mengalami kesulitan dominan pada fonem halkiyah, sedangkan mahasiswa lain lebih banyak melakukan kesalahan pada fonem tafkhim atau interdental. Variasi ini menunjukkan bahwa kesalahan pelafalan tidak hanya dipengaruhi oleh faktor linguistik, tetapi juga oleh pengalaman belajar, intensitas latihan, serta kemampuan individu dalam mengontrol organ artikulasi. Oleh karena itu, pembelajaran Ilmu Ashwat sebaiknya menerapkan pendekatan pembelajaran terdiferensiasi dengan memberikan latihan yang disesuaikan dengan kebutuhan masing-masing mahasiswa berdasarkan hasil diagnosis pelafalan.

Dari sisi pengembangan keilmuan, penelitian ini memperkuat relevansi integrasi antara kajian linguistik Arab dan teknologi kecerdasan buatan dalam pembelajaran bahasa. Selama ini, penelitian Ilmu Ashwat lebih banyak berfokus pada analisis deskriptif mengenai makharijul huruf dan sifat bunyi, sedangkan penelitian di bidang teknologi bahasa berkembang secara terpisah. Melalui penelitian ini diperlihatkan bahwa kedua bidang tersebut dapat saling melengkapi sehingga menghasilkan model evaluasi pelafalan yang lebih objektif, efisien, dan dapat direplikasi pada berbagai institusi pendidikan Islam. Integrasi tersebut membuka peluang penelitian lanjutan mengenai pengembangan sistem evaluasi pelafalan otomatis yang secara khusus dirancang berdasarkan karakteristik fonologi bahasa Arab bagi penutur bahasa Indonesia.

Temuan tersebut juga mengindikasikan bahwa pemetaan kesalahan pelafalan menggunakan speech recognition dapat dimanfaatkan sebagai dasar penyusunan program remedial. Mahasiswa yang mengalami kesalahan pada kelompok fonem tertentu dapat memperoleh latihan yang lebih terfokus dibandingkan mahasiswa yang telah menguasai kelompok fonem tersebut. Pendekatan ini memungkinkan proses pembelajaran menjadi lebih efisien karena dosen dapat mengalokasikan waktu pembelajaran sesuai kebutuhan nyata mahasiswa.

Ditinjau dari perspektif analisis kesalahan (Corder, 1967), sebagian besar kesalahan yang teridentifikasi dalam penelitian ini tergolong kesalahan interlingual, yakni kesalahan yang bersumber dari transfer negatif sistem fonologi bahasa pertama ke dalam pelafalan bahasa target, sebagaimana tampak pada pergantian bunyi faringal menjadi glotal maupun interdental menjadi alveolar. Namun demikian, catatan validator juga mengindikasikan adanya sejumlah kesalahan yang bersifat developmental, yaitu ketidakkonsistenan pelafalan pada fonem yang sama oleh mahasiswa yang sama pada posisi kata yang berbeda, yang lebih mencerminkan keterbatasan otomatisasi motorik artikulasi dibandingkan transfer bahasa pertama semata.

Perbedaan sumber kesalahan ini memiliki implikasi pedagogis yang berbeda: kesalahan interlingual memerlukan penyadaran kontrastif (contrastive awareness raising) atas perbedaan titik artikulasi, sedangkan kesalahan developmental lebih tepat diatasi melalui latihan artikulasi (drilling) yang intensif dan berulang.

Secara keseluruhan, faktor penyebab kesalahan yang teridentifikasi dari hasil wawancara dan catatan validator dapat dikelompokkan menjadi dua: (1) interferensi fonologis bahasa Indonesia sebagai bahasa pertama, yang menyebabkan mahasiswa secara tidak sadar menggantikan bunyi Arab yang tidak dikenal dengan bunyi terdekat dalam bahasa Indonesia; dan (2) minimnya latihan artikulasi makhraj secara khusus dan berulang, sehingga otot-otot artikulasi (lidah, faring, langit-langit) belum terlatih untuk menghasilkan bunyi-bunyi asing tersebut secara otomatis. Temuan ini memperkuat argumen Balula (2021) dan Abdou serta Rashwan (2014) bahwa umpan balik otomatis dan berulang dari sistem ASR berpotensi mempercepat proses pembiasaan artikulasi dibandingkan koreksi manual yang terbatas pada jam tatap muka di kelas.

Secara pedagogis, temuan ini menyarankan agar pembelajaran makharijul huruf pada Program Studi Pendidikan Bahasa Arab tidak lagi diperlakukan sebagai materi pengantar yang cukup diajarkan sekali di awal perkuliahan, melainkan sebagai keterampilan motorik yang memerlukan latihan berulang dan terpantau, khususnya untuk kelompok fonem halkiyah dan tafkhim yang terbukti paling sulit dikuasai. Pemanfaatan aplikasi speech recognition sebagai alat latihan mandiri berpotensi memberikan umpan balik yang lebih sering dibandingkan koreksi tatap muka yang terbatas pada jam perkuliahan, meskipun penggunaannya tetap perlu didampingi evaluasi berkala oleh dosen Ilmu Ashwat untuk mengonfirmasi validitas umpan balik otomatis tersebut, terutama pada kasus pelafalan yang sulit dibedakan secara akustik.

Implikasi praktis penelitian ini tidak hanya ditujukan kepada dosen Ilmu Ashwat, tetapi juga kepada pengelola program studi. Hasil pemetaan kesalahan pelafalan dapat dijadikan dasar dalam mengevaluasi capaian pembelajaran mata kuliah Ilmu Ashwat serta menyusun program penguatan kompetensi makharijul huruf pada semester awal. Selain itu, hasil penelitian ini dapat menjadi referensi dalam pengembangan laboratorium bahasa berbasis teknologi speech recognition sehingga mahasiswa memperoleh kesempatan berlatih secara mandiri dengan tetap memperoleh umpan balik yang terukur.

Penelitian ini tidak terlepas dari sejumlah keterbatasan. Pertama, subjek penelitian hanya mencakup satu angkatan pada satu program studi di satu institusi, sehingga generalisasi temuan ke konteks lain perlu dilakukan secara hati-hati. Kedua, aplikasi speech

recognition yang digunakan pada dasarnya dilatih dengan data bahasa Arab yang didominasi penutur asli, sehingga tingkat akurasinya terhadap variasi pelafalan pembelajar bahasa Arab sebagai bahasa asing berpotensi berbeda dari akurasi yang dilaporkan pada konteks penutur asli. Ketiga, penelitian ini belum menelusuri secara mendalam faktor afektif seperti kecemasan berbicara (speaking anxiety) yang berpotensi turut memengaruhi performa pelafalan mahasiswa saat direkam, sehingga aspek ini dapat menjadi arah penelitian lanjutan

SIMPULAN

Penelitian ini menyimpulkan bahwa mahasiswa Program Studi Pendidikan Bahasa Arab Institut Agama Islam Pemalang paling banyak mengalami kesalahan pelafalan pada fonem-fonem halkiyah, khususnya ع, diikuti oleh fonem tafkhim dan interdental, sementara fonem yang memiliki padanan bunyi dalam bahasa Indonesia relatif jarang salah dilafalkan. Hasil deteksi speech recognition menunjukkan kesesuaian yang tinggi dengan penilaian validator ahli, mengindikasikan bahwa teknologi ASR dapat difungsikan sebagai alat bantu skrining awal yang efisien dalam analisis kesalahan pelafalan bahasa Arab, tanpa menggantikan peran penilaian pakar Ilmu Ashwat sepenuhnya. Penyebab utama kesalahan adalah interferensi fonologis bahasa Indonesia dan minimnya latihan artikulasi makhraj secara intensif.

Kontribusi dan kebaruan penelitian ini terletak pada penggabungan pendekatan speech recognition dengan analisis linguistik Ilmu Ashwat serta validasi pakar dalam satu prosedur diagnosis pelafalan yang sistematis, sesuatu yang belum banyak diterapkan pada konteks pembelajaran bahasa Arab di perguruan tinggi keagamaan Islam di Indonesia. Hasil penelitian ini dapat dimanfaatkan program studi sebagai dasar evaluasi kurikulum pembelajaran Ilmu Ashwat dan makharijul huruf, serta sebagai landasan pengembangan modul latihan artikulasi berbantuan teknologi speech recognition secara mandiri bagi mahasiswa

DAFTAR PUSTAKA

- Abdou, S. M., & Rashwan, M. (2014). A computer aided pronunciation learning system for teaching the Holy Quran recitation rules. In 2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA) (pp. 543–550). IEEE.
- Al-Ani, S. H. (1970). Arabic phonology: An acoustical and physiological investigation. Mouton.
- Al-Marri, M., Raafat, H., Abdallah, M., Abdou, S., & Rashwan, M. (2018). Computer aided Qur'an pronunciation using DNN. *Journal of Intelligent & Fuzzy Systems*, 34(5), 3257–3271.

- Alqadasi, A. M., Zeki, A. M., Sunar, M. S., Salam, M. S. B. H., Abdulghafor, R., & Khaled, N. A. (2023). Improving automatic forced alignment for phoneme segmentation in Quranic recitation. *IEEE Access*, 12, 229–244.
- Alsahafi, Y. S., & Asad, M. (2024). Empirical study on mispronunciation detection for tajweed rules during Quran recitation. In *2024 6th International Conference on Computing and Informatics (ICCI)* (pp. 39–45). IEEE.
- Aziz, M. A., & Masyithoh, S. (2023). Problematika pelafalan huruf hijaiyah pada usia dewasa (perspektif Ilmu Ashwat). *Arabic Language in Focus*, 1(1), 12–18.
- Balula, N. O. (2021). Automatic speech recognition (ASR) systems for learning Arabic language and Al-Quran recitation: A review. *International Journal of Computer Science and Mobile Computing*, 10(7), 91–100.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37–46.
- Corder, S. P. (1967). The significance of learners' errors. *International Review of Applied Linguistics*, 5, 161–170.
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage.
- Dewi, M. I. N. (2021). Analisis kesalahan pronunciation pada mahasiswa tingkat 1 Universitas Kebangsaan. *Jurnal TEDC*, 15(1), 101–106.
- Hafidz bin Zaid, Abdul, dkk. (2024). Implementasi Pendekatan Komunikatif Dalam Meningkatkan Keterampilan Berbahasa Arab. 7(2). 682–688. <http://dx.doi.org/10.35931/am.v7i2.3769>
- Irbabul Lubab, Mohamad. Dkk (2024). Program Shabāhul Lughah Untuk Pembelajaran Mufradāt Siswa Madrasah Aliyah Negeri 1 Jepara. 7(2). <http://dx.doi.org/10.35931/am.v7i2.3977>
- Karimah, I., Letmiros, & Hajidah, G. N. (2022). Interferensi fonologis bahasa Indonesia terhadap bahasa Arab pada pembacaan Surah al-Fātiḥah. *Kalimatuna: Journal of Arabic Research*, 1(1), 81–104. <https://doi.org/10.15408/kjar.v1i1.27198>
- Kheir, Y. E., Ibrahim, O., Meghanani, A., Almarwani, N., Toyin, H. O., Alharbi, S., Alfadly, M., Alkanhal, L., Selim, I., & Elbatal, S. (2025). Towards a unified benchmark for Arabic pronunciation assessment: Quranic recitation as case study (arXiv:2506.07722). arXiv.
- Kosim, N., Ardiyansyah, A. A., & Aryawan, S. (2024). Analisis kesalahan makharijul huruf santri madrasah As-Salam. *Diwan: Jurnal Bahasa dan Sastra Arab*, 16(2), 140–149.
- Lado, R. (1957). *Linguistics across cultures: Applied linguistics for language teachers*. University of Michigan Press.
- Marlina, L. (2019). Analisis kontrastif fonologi bahasa Arab dan bahasa Indonesia dalam pembelajaran pidato bahasa Arab pada Program Studi Pendidikan Bahasa Arab UIN Sunan Gunung Djati Bandung. *Metalingua: Jurnal Penelitian Bahasa*, 18(2), 125–134.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). Sage.
- Nasution, N., & Lubis, L. (2023). Analisis kesalahan makharijul huruf pada pelafalan kalimat bahasa Arab kelas VIII MTs Al-Jam'iyatul Washliyah Tembung. *MODELING: Jurnal Program Studi PGMI*, 10(2), 223–230.
- Nurshafnita, P., & Zainuddin, D. (2023). Analisis kesalahan pelafalan huruf hijaiyah kelas VII MTs Al-Wasliyah Sigambal. *Tsaqofiya: Jurnal Pendidikan Bahasa dan Sastra Arab*, 5(2), 196–213. <https://doi.org/10.21154/tsaqofiya.v5i2.198>
- Rahmatia, Darwis, M., & Lukman. (2021). Analisis kesalahan fonologi dalam keterampilan membaca teks bahasa Arab. *Nady Al-Adab*, 18(1), 121–139. <https://doi.org/10.20956/jna.v18i1.13516>
- Sugiyono. (2022). *Metode penelitian kuantitatif, kualitatif, dan R&D* (Cet. 29). Alfabeta.
- Supriyanoor, dkk. (2025). Metode Eklektik Dalam Pembelajaran Bahasa Arab: Integrasi Strategis untuk Hasil Maksimal. 8(2), 446–453. <http://doi.org/10.35931/am.v8i2.5240>
- Thoyib, T., & Hamidah, H. (2018). Interferensi fonologis bahasa Arab: Analisis kontrastif fonem bahasa Arab terhadap fonem bahasa Indonesia pada mahasiswa Universitas Al Azhar bukan Jurusan Sastra Arab. *Jurnal Al-Azhar Indonesia Seri Humaniora*, 4(2), 63–71. <https://doi.org/10.36722/sh.v4i2.257>
- Zuliyah Safitri, dkk. (2025). Integrating Affective Evaluation in Arabic Language Learning: A Theoretical and Conceptual Analysis within the Indonesian Islamic Education. 8(2). 528–536. <http://doi.org/10.35931/am.v8i2.5170>