

INSTITUTO FEDERAL DE CIÊNCIA E TECNOLOGIA DE SÃO PAULO
CAMPUS BRAGANÇA PAULISTA (IFSP-BRA)

BRUNA CAMARGO GALINDO
FERNANDA NASCIMENTO DE PAULA

**PARLA: PLATAFORMA PARA AUXILIAR NO APRENDIZADO
DE DEFICIENTES AUDITIVOS E SURDOS NA ORALIZAÇÃO
E NA LEITURA LABIAL – FASE II**

Bragança Paulista
2025

BRUNA CAMARGO GALINDO
FERNANDA NASCIMENTO DE PAULA

**PARLA: PLATAFORMA PARA AUXILIAR NO APRENDIZADO
DE DEFICIENTES AUDITIVOS E SURDOS NA ORALIZAÇÃO
E NA LEITURA LABIAL – FASE II**

Relatório solicitado como parte dos requisitos para a apresentação de trabalho de pesquisa científica para a 15ª BRAGANTEC, realizada de 16 a 18 de outubro de 2025, no IFSP *campus* Bragança Paulista.

Orientador: Prof. Dr. Rafael Prearo Lima
Coorientadora: Profa. Me. Talita de P. C. de Souza

Bragança Paulista
2025

RESUMO

O objetivo deste projeto é desenvolver uma plataforma para auxiliar no processo de oralização e de aprendizagem de leitura labial para surdos e para pessoas com surdez moderada ou profunda. Alinhados aos ODS 3 (Saúde e Bem-estar), ODS 4 (Educação de Qualidade) e ODS 10 (Redução das Desigualdades), buscamos facilitar a comunicação entre esse público e o mundo ouvinte. Para isso, criamos um banco de dados de gravações de vozes e ruídos de fundo, treinamos uma IA para reconhecer os sons da fala e desenvolvemos GIFs animados que ilustram, em detalhe, os movimentos faciais e labiais de cada fonema. Na Fase I, desenvolvida em 2024, dedicada às vogais do português brasileiro (A, E, É, I, O, Ó e U), utilizamos mais de 77 000 gravações no treinamento. Nos primeiros testes com ouvintes, o protótipo atingiu 80 % de acertos; após calibrações de sensibilidade, filtragem de ruídos e ampliação da base para cerca de 80 000 amostras, chegou a 90 % e, com ajustes finais, aproximou-se de 100 %. Na Fase II, fase atual, integraram-se ao banco nove consoantes (S, P, B, D, G, Z, T, F e V) e 45 sílabas simples (CA/KA, SA/SE/SI/SO/SU, PA/PE/PI/PO/PU, BA/BE/BI/BO/BU, DA/DE/DI/DO/DU, GA/GE/GI/GO/GU, ZA/ZE/ZI/ZO/ZU, TA/TE/TI/TO/TU, FA/FE/FI/FO/FU e VA/VE/VI/VO/VU). Nos testes iniciais, o índice foi de pouco mais de 60%; após adicionar milhares de registros, subiu a 80%. Para alcançar precisão próxima de 100%, implementamos um indicador gráfico de intensidade de pronúncia em cada nível e passamos a usar o software Praat para análise detalhada.

Palavras-chave: Oralização. Surdos. Leitura labial.

ABSTRACT

The objective of this project is to develop a platform to assist in the process of speech training and lip-reading for deaf individuals and people with moderate or profound hearing loss. Aligned with SDG 3 (Good Health and Well-being), SDG 4 (Quality Education), and SDG 10 (Reduced Inequalities), we aim to facilitate communication between this group and the hearing world. To achieve this, we created a database of voice recordings and background noises, trained an AI to recognize speech sounds, and developed animated GIFs that illustrate, in detail, the facial and lip movements of each phoneme. In Phase I, developed in 2024 and dedicated to the vowels of Brazilian Portuguese (A, E, É, I, O, Ó, and U), we used more than 77,000 recordings in training. In the first tests with hearing individuals, the prototype achieved an accuracy rate of 80%; after sensitivity calibrations, noise filtering, and an expansion of the database to approximately 80,000 samples, it reached 90%, and with final adjustments, it neared 100%. In Phase II, the current phase, nine consonants (S, P, B, D, G, Z, T, F, and V) and 45 simple syllables (CA/KA, SA/SE/SI/SO/SU, PA/PE/PI/PO/PU, BA/BE/BI/BO/BU, DA/DE/DI/DO/DU, GA/GE/GI/GO/GU, ZA/ZE/ZI/ZO/ZU, TA/TE/TI/TO/TU, FA/FE/FI/FO/FU, and VA/VE/VI/VO/VU) have been integrated into the database. Initial tests showed an accuracy rate of just over 60%; after adding thousands of new recordings, it increased to 80%. To achieve near-perfect accuracy, we implemented a graphical pronunciation intensity indicator at each level and began using the Praat software for detailed analysis:

Keywords: Speech training. Deaf individuals. Lip-reading.

AGRADECIMENTOS

À organização da 15ª Bragantec, pela disponibilização do espaço e pela oportunidade de divulgação científica em ambiente propício à inovação e à inclusão.

Aos nossos orientadores, por seu apoio contínuo, cuja orientação técnica e acadêmica foi essencial para o desenvolvimento e aprimoramento do trabalho.

Aos docentes do IFSP-BRA, que contribuíram significativamente com seus conhecimentos e sugestões ao longo do processo.

Aos voluntários, pela valiosa colaboração na construção do banco de dados da inteligência artificial, fornecendo insumos fundamentais para a evolução da plataforma.

Aos nossos familiares, pelo incentivo, suporte emocional e confiança depositada, elementos indispensáveis para a concretização deste projeto.

SUMÁRIO

1	INTRODUÇÃO	7
2	OBJETIVOS	9
3	DESENVOLVIMENTO DO PROJETO	10
4	RESULTADOS	17
5	CONCLUSÃO	19
	REFERÊNCIAS	21

1 INTRODUÇÃO

Dados do IBGE, do ano de 2020, divulgados pela Assembleia Legislativa de São Paulo (Alesp, 2021), constataram que, a cada duzentos bebês, três nascem com surdez total no Brasil. Indo além, quando pensamos no número dos que possuem algum grau de deficiência auditiva acima do moderado ou surdez no Brasil, o total já ultrapassa a marca de 10 milhões de brasileiros (Lemos, 2023).

De acordo com o Censo Escolar do Inep (INEP, 2023), dos 47,3 milhões de alunos da educação básica, 61.594 possuem alguma deficiência relacionada à surdez, o que indica que, talvez, um número grande esteja fora do ambiente escolar. Nesse sentido, um dos desafios para essas pessoas é a questão dos estudos, visto que, por falta de ensino adequado, ou por falhas em sua adaptação, grande parte deles acaba abandonando o ensino formal. Crianças, por exemplo, no ambiente escolar, podem encontrar barreiras que as impeçam de ser incluídas nos estudos, em brincadeiras e até mesmo para que sejam ajudadas por aqueles que não têm conhecimento da Libras. Jovens e adultos, por sua vez, mesmo sabendo Libras, podem perder oportunidades de emprego por obstáculos na comunicação (além de outras dificuldades, como aponta Magno, 2021).

A partir dessa necessidade, descobrimos que várias pessoas surdas conseguiram ser oralizadas. Uma delas foi Alex Bill (do perfil @alexbilljr). Diagnosticado com deficiência auditiva bilateral com um ano de idade, ele é um surdo oralizado que aprendeu a falar, entre outros fatores, com a ajuda da leitura labial e hoje trabalha como palestrante e digital influencer.

O caso de Alex Bill não é isolado. Alguns estudos, como o de Toffolo et al. (2017) e o de Pinheiro et al. (2020), apontam para os benefícios que o desenvolvimento das habilidades de leitura labial e da oralização trazem para a pessoa com surdez. Dentre eles, estão uma melhor comunicação dos surdos com o mundo ouvinte e uma maior facilidade no processo de leitura. Assim, tanto a oralização como a leitura labial podem auxiliar os surdos no ambiente escolar, o que será importante não apenas para seu futuro acadêmico, como também para o profissional.

Com base nesses dados, decidimos criar o Parla, uma plataforma que auxilia na oralização e na aprendizagem de leitura labial de pessoas com deficiência auditiva severa ou profunda e de surdos. Treinado por meio de IA, o Parla serve como uma ferramenta de apoio para esses indivíduos que buscam outras formas, além da Língua Brasileira de Sinais (Libras), de interagir com o mundo ouvinte.

Crucial para o desenvolvimento humano como ser social e intelectual, a comunicação é uma das dificuldades enfrentadas por pessoas com deficiência auditiva severa ou profunda e

por surdos, gerando impactos em diferentes áreas de sua vida, dentre as quais destacamos a educação. Apesar da existência da Libras, parte dessa população não consegue se comunicar com aqueles do mundo ouvinte porque estes, em sua grande maioria, não sabem se comunicar em Libras, quer pela dificuldade de aprendê-la, quer por falhas no sistema de ensino brasileiro, ou até mesmo por desinteresse. Assim, justificamos a criação do Parla como uma ferramenta disponível para que pessoas com deficiência auditiva severa ou profunda e também os surdos possam, caso queiram, ser oralizadas e aprendam a fazer leitura labial, a fim de facilitar sua comunicação com o mundo ouvinte.

Também justificamos a criação dessa plataforma a partir de três dos dezessete Objetivos de Desenvolvimento Sustentável (ODS) recomendados pela ONU (2024), a saber, os objetivos (3) Saúde e Bem-estar, entendendo que a facilidade na comunicação pode possibilitar novos relacionamentos, o que traz benefícios para a saúde mental dos indivíduos; (4) Educação de Qualidade, visto que a plataforma pode facilitar a comunicação no ambiente educacional, tornando-o mais inclusivo, o que pode garantir que todos os estudantes, independentemente de suas capacidades, tenham acesso às mesmas oportunidades de aprendizagem e desenvolvimento; e (10) Redução das desigualdades, pois promove a inclusão social e econômica de surdos ao mitigar barreiras de comunicação, permitindo que esses indivíduos tenham acesso igualitário a oportunidades e serviços.

Diante desse cenário, levantamos as seguintes perguntas-problema, que direcionam nosso projeto: como podemos auxiliar surdos e pessoas com deficiência auditiva severa ou profunda em sua comunicação com o mundo ouvinte? Como podemos treinar uma IA para ajudar na oralização e na aprendizagem de leitura labial desses indivíduos?

Nossa hipótese é a de que, por meio do treino de uma IA para reconhecer os sons de fala, e com a ajuda de gifs com indicação de movimento faciais (dos lábios, da boca, da língua), o Parla possa auxiliar pessoas com deficiência auditiva severa ou profunda e surdos no processo de oralização e de aprendizagem de leitura labial.

2 OBJETIVOS

2.1 Objetivo geral

O objetivo geral do Parla é desenvolver uma plataforma para auxiliar no processo de oralização e de aprendizagem de leitura labial para surdos e para pessoas com deficiência auditiva severa ou profunda.

2.1 Objetivos específicos

Com base nesse objetivo geral, estabelecemos os seguintes objetivos específicos divididos em fases de desenvolvimento:

- *Fase I* (desenvolvida em 2024): (i) desenvolver uma plataforma online, dividida em níveis, fases e etapas; (ii) montar um banco de dados composto por gravação de vozes de vogais; (iii) começar a treinar uma IA, a partir de um banco de dados com vozes, para reconhecer os sons produzidos pela fala humana, e criar gifs com movimentos faciais e labiais para ensinar os usuários como os sons são produzidos.
- *Fase II* (em desenvolvimento – 2025): (i) implementar ao banco de dados as gravações das consoantes e sílabas simples; (ii) continuar o treinamento da IA, a partir de um banco de dados com vozes;
- *Fase III* (a ser desenvolvida – 2026): (i) implementar ao banco de dados as gravações de sílabas complexas; (ii) finalizar o treinamento da IA, a partir de um banco de dados com vozes; (iii) viabilizar a operacionalização comercial da plataforma.

3 DESENVOLVIMENTO DO PROJETO

Primeiramente, durante a *Fase I* (desenvolvida em 2024) do projeto, identificamos uma necessidade específica na sociedade que gostaríamos de suprir. Alinhados aos Objetivos de Desenvolvimento Sustentável (ODS), decidimos focar no auxílio a pessoas com deficiência auditiva severa ou profunda, incluindo surdos. Em seguida, elaboramos perguntas-problema para direcionar nosso trabalho e, a partir delas, estabelecemos os objetivos a serem alcançados.

Posteriormente, aprofundamos nossas pesquisas para coletar informações relevantes e compreender como poderíamos apoiar esse público. Após mapear desafios, dificuldades e métodos já existentes, projetamos uma plataforma voltada à oralização e ao ensino de leitura labial para surdos e pessoas com deficiência auditiva severa ou profunda.

Iniciamos o desenvolvimento do protótipo da plataforma Parla criando nossa base de dados. Para gerenciar os cadastros, utilizamos o MySQL. Em seguida, desenvolvemos as telas de evolução, contato, atividades e a tela inicial. Cada uma dessas telas foi projetada considerando a psicologia das cores, que demonstra como cada cor pode influenciar a percepção e o comportamento dos usuários, especialmente crianças, em relação a determinados estímulos.

Por isso, a cor predominante do Parla é o verde. Na psicologia das cores e no contexto da alfabetização, o verde está associado ao equilíbrio e à concentração. Ele promove um ambiente de calma, harmonia e bem-estar, fatores essenciais para um aprendizado eficaz. Além disso, é uma cor que estimula a criatividade, melhora a memória e contribui para a estabilidade do sistema nervoso, sendo, portanto, ideal para ambientes educacionais como nossa plataforma. O verde também favorece a concentração e promove uma visão positiva sobre o site.

Utilizamos HTML para o desenvolvimento do corpo do protótipo, CSS para a estilização, posicionamento e dimensionamento dos elementos, e JavaScript para tornar as páginas web interativas e dinâmicas, possibilitando funções como validação de formulários, integração com a IA, além de outras funcionalidades.

Após consolidarmos a estrutura do protótipo, definimos os níveis do Parla, organizados em fases com conteúdo de dificuldade progressiva: (a) vogais; (b) consoantes; (c) sílabas simples; (d) sílabas complexas. Cada atividade inclui uma imagem GIF com movimentos faciais e, ao lado, a imagem da letra correspondente para o usuário reproduzir.

Utilizamos o Teachable Machine, uma ferramenta essencial para o desenvolvimento e treinamento da IA do projeto, para que a IA, desenvolvida com técnicas de machine learning, possa reconhecer os sons emitidos pelo usuário e filtrar ruídos indesejados com precisão.

Para garantir um funcionamento satisfatório, começamos gravando diversos tipos de ruídos de fundo que poderiam interferir na gravação do usuário em situações normais, como ruído branco, chiados, respirações, conversas paralelas, sons de animais de estimação, materiais escolares, buzinas, sons de cozinha, sons de banheiro, sons de tecidos, entre outros.

Além disso, contamos com a colaboração de várias pessoas para gravar os níveis iniciais com diferentes formas, entonações e tipos de voz, a fim de construir um banco de dados extenso que permitisse à IA realizar a correção da maneira mais eficaz possível.

Ao concluir a base do protótipo, elaboramos uma identidade visual para nossa plataforma, incluindo a criação da mascote Parlinha, um sapo pingo-de-ouro (*Brachycephalus ephippium*), que produz sons, mas não é capaz de ouvir. Essa mascote atua como assistente do usuário na plataforma.

O Parla foi desenvolvido utilizando HTML, CSS e JavaScript, estando integrado a um banco de dados no MySQL Workbench. A plataforma conta atualmente com 71 páginas desenvolvidas, incluindo a tela de apresentação do site, a tela de cadastro, a tela de login, a tela de recuperação de senha, a tela de desenvolvimento do personagem interativo (demonstrativa), a tela inicial do usuário, a tela de fases (que dá acesso a três níveis: básico, intermediário e avançado), com atualmente 35 telas de níveis disponíveis para teste na primeira fase. Na segunda fase, temos 11 conjuntos com 7 sílabas cada, além de a tela de falha, a tela de evolução (demonstrativa) e a tela de contato (demonstrativa). Essa estrutura garante a organização e o progresso do usuário de forma clara e didática, alinhada ao objetivo principal da plataforma: desenvolver habilidades de leitura labial e oralização de forma interativa.

A plataforma é estruturada em três fases: a Base (inicial), o Caule (intermediário) e a Flor (avançado). Cada fase é subdividida em níveis e etapas específicas. Até o momento, as fases foram desenvolvidas da seguinte forma:

A Fase Base envolve o aprendizado inicial das vogais (A, E, É, I, O, Ó, U), que são essenciais para a construção das primeiras palavras. Além disso, os alunos aprendem consoantes e sílabas simples, como Ba, Be, Bi, Bo, Bu, Ca, Ce, Ci, Co, Cu, entre outras, cobrindo uma ampla variedade de combinações fonéticas. Essa etapa é fundamental para que o aluno compreenda a estrutura básica das sílabas e inicie a formação de palavras simples.

Posteriormente, são introduzidas consoantes e sílabas complexas, como Cha, Che, Chi, Cho, Chu, Lha, Lhe, Lhi, Lho, Lhu, além de combinações como Gua, Gue, Gui, Guo. Essas combinações ampliam o repertório linguístico do aluno, permitindo que ele entenda e pronuncie palavras mais elaboradas. A fase também abrange sílabas com encontros consonantais, como Bra, Bre, Bri, Bro, Bru, e sílabas nasais, como Ar, Er, Ir, Or, Ur, An, En, In, On, Un.

Para complementar o aprendizado, a Fase Base inclui outras combinações e variações de sílabas que ajudam a diversificar o conhecimento fonético do aluno. Exemplos dessas combinações são Ça, Ço, Çu, e encontros consonantais mais avançados, como Psa, Pse, Psi, Psu, e outras formas como Tcha, Tche, Tchi, Tcho, Tchu, que permitem um domínio mais completo das estruturas fonológicas do idioma.

Na Fase Caule, o conteúdo aprendido na Fase Base é ampliado, permitindo que os alunos formem várias palavras. Essas palavras são classificadas em diferentes categorias gramaticais, como palavras simples e compostas, substantivos, pronomes, verbos, preposições, artigos, adjetivos, conjunções, interjeições, numerais e advérbios. O objetivo é garantir que os alunos compreendam e utilizem essas classes de palavras de forma adequada, consolidando seu conhecimento gramatical.

Na Fase Flor, o aprendizado é levado um passo adiante. O que foi assimilado na Fase Caule é aplicado para a formação de frases, orações ou sentenças completas. Essa etapa foca na construção de estruturas linguísticas mais complexas, permitindo que os alunos utilizem seu conhecimento de palavras para se expressar de maneira clara e coesa.

Para o desenvolvimento das fases durante o treinamento da plataforma, incluímos uma ampla gama de sílabas e combinações fonéticas, abrangendo:

- Vogais: A, E, É, I, O, Ó, U;
- Consoantes e sílabas simples: Ba, Be, Bi, Bo, Bu, Ca, Ce, Ci, Co, Cu, Da, De, Di, Do, Du, Fa, Fe, Fi, Fo, Fu, Ga, Ge, Gi, Go, Gu, Ha, He, Hi, Ho, Hu, Ja, Je, Ji, Jo, Ju, Ka, Ke, Ki, Ko, Ku, La, Le, Li, Lo, Lu, Ma, Me, Mi, Mo, Mu, Na, Ne, Ni, No, Nu, Pa, Pe, Pi, Po, Pu, Qua, Que, Qui, Quo, Ra, Re, Ri, Ro, Ru, Sa, Se, Si, So, Su, Ta, Te, Ti, To, Tu, Va, Ve, Vi, Vo, Vu, Wa, We, Wi, Wo, Wu, Xa, Xe, Xi, Xo, Xu, Ya, Ye, Yi, Yo, Yu, Za, Ze, Zi, Zo, Zu;
- Consoantes e sílabas complexas: Cha, Che, Chi, Cho, Chu, Lha, Lhe, Lhi, Lho, Lhu, Nha, Nhe, Nhi, Nho, Nhu, Gua, Gue, Gui, Guo, Bla, Ble, Bli, Blo, Blu, Cla, Cle, Cli, Clo, Clu, Fla, Fle, Fli, Flo, Flu, Gla, Gle, Gli, Glo, Glu, Pla, Ple, Pli, Plo, Plu, Bra, Bre, Bri, Bro, Bru, Cra, Cre, Cri, Cro, Cru, Dra, Dre, Dri, Dro, Dru, Fra, Fre, Fri, Fro, Fru, Gra, Gre, Gri, Gro, Gru, Pra, Pre, Pri, Pro, Pru, Tra, Tre, Tri, Tro, Tru, Vra, Vre, Vri, Vro, Vru, Rra, Rre, Rri, Rro, Rru, Ar, Er, Ir, Or, Ur, Al, El, Il, Ol, Ul, An, En, In, On, Un, As, Es, Is, Os, Us, At, Et, It, Ot, Ut, Am, Em, Im, Om, Um, Az, Ez, Iz, Oz, Uz;
- Outras combinações e variações: Ça, Ço, Çu, Bar, Ber, Bir, Bor, Bur, Car, Cer, Cir, Cor, Cur, Dar, Der, Dir, Dor, Dur, Far, Fer, Fir, For, Fur, Gar, Ger, Gir, Gor, Gur, Har, Her, Hir, Hor, Hur, Jar, Jer, Jir, Jor, Jur, Lar, Ler, Lir, Lor, Lur, Mar, Mer, Mir, Mor, Mur, Nar, Ner,

Nir, Nor, Nur, Par, Per, Pir, Por, Pur, Quar, Quer, Quir, Quor, Rar, Rer, Rir, Ror, Rur, Sar, Ser, Sir, Sor, Sur, Tar, Ter, Tir, Tor, Tur, Var, Ver, Vir, Vor, Vur, Xar, Xer, Xir, Xor, Xur, Zar, Zer, Zir, Zor, Zur, Psa, Pse, Psi, Psu, Peso, Ac, Ec, Ic, Oc, Uc, Pna, Pne, Pni, Pno, Pneu, Abs, Pt, Etc, Tsa, Tse, Tsi, Tso, Tsu, Tcha, Tche, Tchi, Tcho, Tchu, É, Ó, ã, ão.

A tela de atividades oferece diversos recursos para apoiar a pronúncia do usuário. Inclui um GIF com movimentos orofaciais que auxilia na aprendizagem de leitura labial, uma imagem demonstrativa da onda sonora que ajuda na entonação e uma caixa de hertz que aparece após a pronúncia do usuário, funcionando como feedback sobre a pronúncia.

Além disso, a tela exibe a imagem do conteúdo que está sendo aprendido no nível, seja letra, sílabas, palavras ou frases. Com base nisso, o usuário tenta pronunciar, e nossa IA captura e analisa o áudio em questão de segundos. Se a pronúncia estiver correta, o usuário avança de fase; caso contrário, é encaminhado para a tela de análise de erros para tentar novamente.

É importante notar que tanto os erros quanto os acertos são salvos no banco de dados para uso futuro na tela de evolução, onde são apresentados os dados do usuário e indicadas atividades extras para as etapas em que o usuário teve maior dificuldade.

A tela de evolução armazena a frequência das atividades, os dias em que foram realizadas, o número de erros e acertos, e a quantidade de tentativas. Com essas informações, é gerado um gráfico de aprendizado do aluno e um boletim que pode ser impresso ou enviado por e-mail.

Na tela de evolução, também são oferecidas atividades extras para recuperação nas áreas em que o usuário apresentou maior número de erros, ajudando a melhorar os pontos de maior dificuldade. O boletim e as atividades de recuperação são disponibilizados mensalmente, permitindo o acompanhamento dos pais, seja quando o aplicativo é utilizado em escolas ou para uso pessoal.

Na Fase II (atualmente em desenvolvimento), integramos ao banco de dados de *machine learning* todas as consoantes e sílabas simples, concluindo assim as Fases I e II do projeto. Nessa etapa, priorizamos os fonemas com maior índice de erro – S, P, B, D, G, Z, T, F e V – além das sílabas:

- CA/KA, SA/SE/SI/SO/SU, PA/PE/PI/PO/PU,
- BA/BE/BI/BO/BU, DA/DE/DI/DO/DU, GA/GE/GI/GO/GU,
- ZA/ZE/ZI/ZO/ZU, TA/TE/TI/TO/TU, FA/FE/FI/FO/FU e VA/VE/VI/VO/VU.

No total, reunimos nove consoantes e quarenta e cinco sílabas sonoras, totalizando 81.080 áudios, correspondendo a 45 horas e 5 minutos de gravação, assegurando dessa forma uma ampla diversidade de modos de fala.

Durante os primeiros testes de desempenho, observamos que as consoantes fricativas (S/Z, V/F, C/K/G, T/D e P/B/D) apresentavam uma taxa de acerto inferior ao esperado – cerca de 60% de acertos em 100 avaliações – devido à variação de entonação entre diferentes falantes. Para aumentar a acurácia, expandimos significativamente o banco de áudios, o que nos permitiu atingir uma taxa de acerto de 80%.

Ainda assim, nosso objetivo era alcançar resultados superiores. Nesse sentido, incorporamos indicadores de intensidade de pronúncia em cada nível e passamos a utilizar o software livre Praat, que nos permitiu realizar análises detalhadas de parâmetros acústicos, como frequência, intensidade, duração, formantes, pitch e qualidade vocal. A partir da visualização das formas de onda, conseguimos otimizar o treinamento do modelo, elevando a precisão da IA para quase 100%.

Para a Fase III (prevista para iniciar em 2026), planejamos incluir no treinamento da inteligência artificial (IA) as sílabas complexas e os encontros consonantais — como BRA, PLA e TRE – utilizando um novo banco de dados composto por vozes humanas de diferentes faixas etárias e sotaques regionais. Essa diversidade sonora visa ampliar ainda mais a capacidade de reconhecimento e compreensão da IA, garantindo um atendimento mais preciso e inclusivo para usuários de diferentes origens linguísticas e culturais.

Paralelamente, estamos desenvolvendo GIFs educativos em alta resolução, com movimentos faciais e labiais detalhados, além de vídeos curtos legendados e animações interativas que ilustram visualmente a articulação de cada som, auxiliando o usuário na reprodução precisa da pronúncia.

Além disso, realizaremos testes de usabilidade em escolas e centros de reabilitação auditiva para coletar *feedback* direto dos usuários, ajustar os níveis de dificuldade e aprimorar a efetividade pedagógica da plataforma. Pretendemos também firmar parcerias estratégicas com instituições de pesquisa, empresas de tecnologia assistiva e órgãos governamentais, viabilizando a implantação comercial do Parla. Nessa etapa, definiremos modelos de licenciamento educacional, planos de manutenção e suporte técnico, bem como estratégias de divulgação para garantir a ampla adoção da plataforma.

No médio prazo, planejamos expandir significativamente as telas atualmente demonstrativas e adicionar mais seis novas telas para tornar o site ainda mais completo e funcional. Essas novas telas incluem: pacotes Premium, atividades Premium, perfil do usuário, recados/mensagens do sistema e uma seção de medalhas e conquistas. É importante destacar que a tela de recuperação de senha já foi implementada. Todas essas interfaces contarão com

melhorias no design, animações e a integração do personagem interativo Parlinha, reforçando a identidade da plataforma e o engajamento do usuário.

Além disso, nosso cronograma prevê a finalização da primeira fase de níveis e o avanço para a segunda fase, que está em andamento. Nesta etapa, estamos focando especialmente nas sílabas com maior índice de erro, como:

- Ca, Ce, Ci, Co, Cu
- Ke, Ki
- Sa, Se, Si, So, Su
- Pa, Pe, Pi, Po, Pu
- Ba, Be, Bi, Bo, Bu
- Da, De, Di, Do, Du
- Ga, Ge, Gi, Go, Gu
- Za, Ze, Zi, Zo, Zu
- Tá, Te, Ti, To, Tu
- Fã, Fé, Fi, Fo, Fu
- Vá, Ve, Vi, Vó, Vu

Esse refinamento visa melhorar a acurácia da IA e apoiar o desenvolvimento de habilidades fundamentais de leitura labial e oralização.

Após estudos aprofundados em gamificação, pretendemos incorporar atividades variadas e motivadoras, inspiradas em plataformas como Duolingo e jogos de alfabetização infantil, para apoiar os focos principais do Parla: a leitura labial e a oralização.

Outro aspecto fundamental que abordaremos é a segurança dos dados. Pretendemos implementar sistemas de segurança robustos, backups de informações e comunicação segura com o servidor, utilizando JavaScript e e-mail. Essa preocupação é especialmente relevante, pois lidamos com dados de crianças e adolescentes, exigindo um sistema atualizado e em conformidade com as exigências do mercado digital. Estamos considerando adotar as diretrizes do OWASP (Open Web Application Security Project), que fornece recomendações de segurança para aplicações web.

Entre os benefícios do uso do OWASP estão: proteção contra ataques comuns (como injeção de SQL), autenticação reforçada, criptografia de dados, monitoramento constante e resposta a incidentes de segurança, além da conformidade com regulamentações de privacidade, como a GDPR e a LGPD. Também pretendemos implementar firewalls e sistemas de detecção de intrusos, para proteger os dados dos usuários e garantir o funcionamento seguro e estável da plataforma.

Em relação ao sistema de backup, será necessário manter três bancos de dados distintos: um para o sistema de login, outro para usuários gratuitos e um para usuários Premium. Para isso, utilizaremos bancos de dados NoSQL, que permitem armazenar dados em diferentes formatos e atualizá-los de acordo com as necessidades de expansão e manutenção, facilitando a escalabilidade do sistema.

Por fim, planejamos adotar um modelo de monetização híbrido, com uma versão gratuita financiada por anúncios e uma versão Premium, cujas receitas ajudarão a custear hospedagem, desenvolvimento, manutenção e suporte contínuo, garantindo a atualização e evolução constante da plataforma.

4 RESULTADOS

Na Fase I, dedicada ao ensino das vogais (A, E, É, I, O, Ó e U), utilizamos um banco de dados com mais de 77 mil gravações para treinar a inteligência artificial (IA). Nos primeiros testes realizados com usuários ouvintes, o protótipo apresentou uma taxa de acertos de 80%, com os pares É/E e Ó/O sendo os principais responsáveis pelos erros de reconhecimento, devido à semelhança de pronúncia entre essas vogais. Para superar essas dificuldades, realizamos ajustes e calibrações no modelo, incluindo a calibração de sensibilidade, uma filtragem refinada de ruídos de fundo e a ampliação da base de treinamento. Essas melhorias elevaram o índice de acertos para 90% e, com um último conjunto de otimizações, alcançamos valores próximos de 100% de precisão. Após a conclusão dessa etapa, o banco de dados foi expandido com novas amostras de vogais e ruídos ambientais, totalizando cerca de 80 mil gravações.

Na Fase II, que está em andamento atualmente, passamos a integrar consoantes (S, P, B, D, G, Z, T, F e V) e sílabas simples. Nos primeiros testes, o protótipo atingiu uma taxa de acertos de aproximadamente 60%. Para melhorar esse desempenho, adicionamos áudios suplementares à base de dados, elevando o índice de acertos para 80%. Visando superar esse patamar, implementamos um indicador gráfico de intensidade da pronúncia em cada nível e incorporamos o uso do software Praat para análises detalhadas dos parâmetros de voz. Essas estratégias permitiram alcançar uma taxa de precisão próxima de 100%, de acordo com as métricas internas de validação.

Até o momento, desenvolvemos integralmente a primeira fase da plataforma Parla, que corresponde ao ensino das vogais. Essa etapa foi fundamental para validar o protótipo e identificar os principais desafios no reconhecimento auditivo. Os erros mais frequentes ocorreram no reconhecimento dos fonemas É e Ó, frequentemente confundidos com E e O. Após ajustes na inteligência artificial, o desempenho do protótipo foi significativamente aprimorado, chegando a resultados muito próximos de 100% de precisão.

Para as próximas fases de testes, buscamos autorização para realizar experimentos com pessoas surdas e com deficiência auditiva severa ou profunda, de modo a validar a eficiência da IA com o público-alvo. Nas etapas seguintes, planejamos expandir o treinamento da IA para lidar com as dependências fonológicas do português brasileiro e, posteriormente, para incluir sílabas simples e complexas.

Após a finalização das gravações específicas para a IA, realizaremos testes com voluntários que não participaram das etapas de gravação, a fim de avaliar o desempenho do sistema em situações reais e corrigir eventuais falhas. Com base nesses testes e em novas gravações direcionadas, pretendemos elevar ainda mais a precisão do reconhecimento.

Na Fase III, atualmente em fase de planejamento, solicitaremos a aprovação para a testagem com pessoas surdas e com deficiência auditiva severa ou profunda, a fim de enriquecer o banco de dados com vozes reais do público-alvo e identificar ajustes específicos de usabilidade. Paralelamente, concluiremos o treinamento da IA com sílabas complexas e encontros consonantais (como “pla”, “tre” e “bra”), além de finalizar o desenvolvimento das telas de perfil do usuário, o mural de recados e as animações educativas em GIF. Esses recursos visam não apenas consolidar o reconhecimento preciso dos sons, mas também oferecer suporte visual e interativo contínuo, preparando o Parla para sua futura operação comercial.

5 CONCLUSÃO

O projeto Parla foi idealizado com o objetivo de criar uma plataforma inovadora e inclusiva, destinada a auxiliar no processo de oralização e no ensino da leitura labial para pessoas surdas e com deficiência auditiva severa ou profunda. O relatório técnico mostrou que a plataforma, ao integrar inteligência artificial (IA) com recursos visuais interativos, tem potencial para gerar avanços significativos na comunicação e na aprendizagem dos usuários.

Durante o desenvolvimento do Parla, a estrutura da plataforma foi cuidadosamente organizada em fases e etapas, apoiando-se em um banco de dados robusto, composto por mais de 80 mil gravações de sons de vogais (A, E, É, I, O, Ó, U), consoantes (S, P, B, D, G, Z, T, F e V) e sílabas simples e complexas (por exemplo, “pla”, “tre”, “bra”). A IA foi treinada com esse acervo sonoro para aprimorar o reconhecimento de sons produzidos pela fala humana. Na Fase I, dedicada às vogais, os primeiros testes com usuários ouvintes revelaram uma taxa de acerto inicial de 80%, com erros concentrados nos pares É/E e Ó/O devido à semelhança de pronúncia. Após calibrações na sensibilidade da IA, filtragem de ruídos de fundo e ampliação da base de treinamento, a precisão subiu para 90%. Com otimizações adicionais, a plataforma alcançou resultados próximos de 100% de acertos.

Na Fase II, em andamento, foram incluídas as consoantes e sílabas simples. Inicialmente, a taxa de acertos foi de cerca de 60%, mas após ajustes (como a inclusão de indicadores gráficos de intensidade da pronúncia e integração do software Praat para análise detalhada dos parâmetros de voz), a IA passou a atingir níveis de precisão similares aos da Fase I, aproximando-se de 100%.

Além do treinamento da IA, o Parla incorporou GIFs com animações faciais e labiais para ensinar a produção correta dos sons, ampliando a interatividade e a eficácia pedagógica. Essa combinação de inteligência artificial com recursos visuais fortaleceu o potencial da plataforma para impactar positivamente a aprendizagem e a autonomia dos usuários.

A análise dos resultados obtidos até agora indica que o Parla poderá contribuir de forma significativa para a inclusão educacional e social de pessoas com deficiência auditiva, favorecendo a comunicação com o público ouvinte e promovendo uma melhor integração nos contextos escolar e profissional. Contudo, reconhecemos algumas limitações no estudo, como a ausência de testes formais com o público-alvo (pessoas surdas e com deficiência auditiva severa ou profunda) e a necessidade de validação contínua da eficácia a longo prazo.

Dadas as dimensões do projeto, e considerando os desafios técnicos e pedagógicos envolvidos, embora saibamos que há surdos com deficiências múltiplas, esse não será o público-alvo neste momento de desenvolvimento da plataforma. A escolha por delimitar o escopo tem como objetivo garantir maior precisão nos testes e na construção das funcionalidades, sem descartar a possibilidade de futuras adaptações que contemplem esse grupo.

Diante desses resultados, planejamos concluir a expansão da plataforma, incluindo recursos visuais e interativos adicionais, bem como realizar testes com um público mais diversificado. A integração do Parla com programas educacionais existentes poderá ampliar seu impacto, garantindo que mais pessoas se beneficiem da tecnologia. Além disso, estamos desenvolvendo novas funcionalidades para atender às demandas emergentes dos usuários.

Para futuras pesquisas, pretendemos investigar como a plataforma pode ser adaptada para diferentes contextos culturais e educacionais, além de avaliar a integração com tecnologias emergentes. Também reconhecemos a importância da segurança dos dados e da manutenção de um sistema robusto: implementaremos políticas rigorosas de segurança da informação e backups eficientes para garantir a integridade dos dados e a operação estável da plataforma.

Em suma, o Parla representa um avanço significativo na promoção da inclusão de pessoas com deficiência auditiva, oferecendo uma ferramenta eficaz para comunicação e aprendizagem. Acreditamos que os próximos passos no desenvolvimento do Parla trarão ainda mais benefícios e oportunidades aos usuários, fortalecendo o vínculo entre o público-alvo e o público ouvinte.

REFERÊNCIAS

ALESP. Dia Internacional da Linguagem de Sinais procura promover a inclusão de pessoas surdas. 23 set. 2021. Disponível em: <https://www.al.sp.gov.br/noticia/?23/09/2021/dia-internacional-da-linguagem-de-sinais-procura-promover-a-inclusao-de-pessoas-surdas->. Acesso em: 09 jun. 2025.

INEP. Confira o panorama dos surdos na educação brasileira. 26 ago. 2023. Disponível em: <https://www.gov.br/inep/pt-br/assuntos/noticias/censo-escolar/confira-o-panorama-dos-surdos-na-educacao-brasileira>. Acesso em: 09 jun. 2025.

LEMONS, Simone. Mais de 10 milhões de brasileiros apresentam algum grau de surdez. **Jornal da USP**. 21 ago. 2023. Disponível em: <https://jornal.usp.br/?p=674626>. Acesso em: 09 jun. 2025.

MAGNO, Rodrigo. As dificuldades da pessoa surda na sociedade brasileira. **Jusbrasil**. 2021. Disponível em: <https://www.jusbrasil.com.br/artigos/as-dificuldades-da-pessoa-surda-na-sociedade-brasileira/1176514129>. Acesso em: 09 jun. 2025.

ONU. Sobre o nosso trabalho para alcançar os Objetivos de Desenvolvimento Sustentável no Brasil. 2024. Disponível em: <https://brasil.un.org/pt-br/sdgs>. Acesso em: 09 jun. 2025.

PINHEIRO, Ângela Maria Vieira et al. Reading strategies for the profoundly deaf Libras users: benefits of speech and lip reading for strengthening linguistic skills. **Estudos de Psicologia**. Campinas, v. 37, 2020. Disponível em: HYPERLINK "<http://dx.doi.org/10.1590/1982-0275202037e190003>"<http://dx.doi.org/10.1590/1982-0275202037e190003>. Acesso em: 09 jun. 2025.

TOFFOLO, Andreia Chagas Rocha; BERNARDINO, Elidéa Lúcia Almeida; VILHENA, Douglas de Araújo; PINHEIRO, Ângela Maria Vieira. Os benefícios da oralização e da leitura labial no desempenho de leitura de surdos profundos usuários da Libras. **Revista Brasileira de Educação**, Rio de Janeiro, v. 22, p. 1-24, 2017. Disponível em: <http://www.scielo.br/scielo.php?pid=S1413-24782017000400227>. Acesso em: 12 ago. 2025.