

O RECONHECIMENTO DE VOZ COMO ALTERNATIVA PARA O ENSINO DA PRONÚNCIA DA LÍNGUA INGLESA EM AULAS REMOTAS: ESTUDO DE CASO COM O APLICATIVO *ELSA SPEAK*

Fábio José de Abreu Moura – (UFPE)
fabiojosedabreumoura@hotmail.com

Diana Vasconcelos Lopes – (UFRPE)
didilaster@gmail.com

RESUMO:

Este trabalho discute o uso do reconhecimento de voz, através do aplicativo ELSA SPEAK, para a prática da pronúncia da língua inglesa (LI), visando incentivar a aprendizagem autônoma dos alunos. O projeto contou com a participação de dois estudantes do ensino fundamental. Para tanto, analisamos as respostas dos alunos a um questionário que investigou o processo de aprendizagem da pronúncia da LI, por meio desse aplicativo.

PALAVRAS-CHAVE: Língua Inglesa. Reconhecimento de Voz. ELSA Speak.

ABSTRACT: This paper discusses the use of voice recognition, through ELSA Speak application, in the practice of English language pronunciation, aiming at stimulating students' autonomous learning. The project counted on the participation of two pupils in basic education. For this purpose, we analyzed their replies to a questionnaire which investigated English pronunciation learning process, with the aid of the aforementioned application.

KEYWORDS: English Language. Voice Recognition. ELSA Speak.

0. Introdução

O uso das Tecnologias Digitais de Informação e Comunicação (doravante TDIC) no ensino é tema de pesquisa e debates há muito tempo. Práticas e métodos são propostos corriqueiramente. Entretanto, quando surge o momento que exige a prática do ensino remoto, percebe-se um cenário educacional, onde prevalece o despreparo de muitos docentes e discentes para o uso adequado da tecnologia a favor da educação, de forma integral.

Forçados a repensar alternativas, um novo momento de ensino floresce e professores e alunos são testados quanto ao letramento digital (esperando que se tenha tal domínio nessa nova ordem da era tecnológica). Vale ressaltar que muitos atores da educação foram surpreendidos com a imediata necessidade de mudança nas práticas de ensino-aprendizagem, até então costumeiramente adotadas.

Como caracteriza Cani et al (2020), “a pandemia acelerou um processo que já estava em curso: a integração das TDIC com a educação”. Isso pode explicar a resposta “uma sensação de despreparo” dos professores da pesquisa exploratória de Honorato e Marcelino (2020). Os dados da pesquisa refletem, de forma generalizada, o pensamento docente quanto aos desafios do ensino remoto. Destarte, comentários como: “falta de acessibilidade de programas e situação econômica dos alunos”; “alunos e professores despreparados para o ensino a distância”; “menor participação nos debates”; “as avaliações não são tão eficazes em medir o aprendizado, ou “os alunos têm um *freepass* para colar” são repetitivos, o que sinaliza para um contexto de dificuldades e os obstáculos impostos a todos os que passaram a participar de situações reais das aulas remotas.

O ensino remoto é de fato um grande desafio; o planejamento é completamente diferente daquele proposto para as aulas presenciais. Com isso, há uma forte mudança na forma como se passou a estabelecer o monitoramento, a fala, a discussão, o componente emocional, o ambiente, o *input*, o *output*, dentre outros fatores que integram o contexto de ensino-aprendizagem.

Importa salientar ainda a resistência de alguns profissionais em incluir essa nova realidade em suas práticas didático-pedagógicas. Ou seja, a necessidade da utilização das TDIC constituiu-se mais como problema do que solução, o que resultou na manutenção da prioridade de práticas tradicionais (HANSEN, 2012).

Pacheco e Lopes (2019) discutem que há uma resistência por parte dos professores ao uso das TDICs. A tais adversidades relaciona-se a não proficiência tecnológica e a pouca prática como estimulante de aprendizagem nas aulas presenciais. Mena (2016) considera urgente que os professores que demonstram alguma forma de resistência “precisam de formação, cursos, infraestrutura, alguém como um especialista em Tecnologias Digitais de Informação e Comunicação (TDIC’s) efetivados dentro da escola” (MENA, 2016, p. 29), isso porque acreditar no digital pode fazer a diferença no aprendizado deles e de seus alunos. Mas, para isso, precisam se encantar pelas tecnologias e suas possibilidades.

Neste sentido, e em face à situação emergencial causada pela pandemia do Covid-19, tornou-se ainda mais urgente a procura por novos métodos de ensinar e aprender, considerando o que se tem disponível em casa, ou mesmo o relacionamento familiar e a vida de distanciamento social. Como afirma Honorato e Marcelino (2020) “é preciso explorar novas maneiras de aprendizagem e interações que estimulem a curiosidade e a criatividade nos alunos para além do currículo acadêmico” (HONORATO; MARCELINO, 2020: 211). Logo, parece plausível dizer que o papel docente está em influenciar o aluno a utilizar os materiais dispostos e compreendê-los de forma autodidata, ao invés de receber esse conhecimento diretamente e na íntegra, apenas através do professor (FREIRE, 2008).

Quando falamos do ensino de pronúncia de língua estrangeira (doravante LE) em contexto remoto, a imprecisão dos sons pode ser um forte empecilho, visto que fatores diversos certamente interferem na percepção e avaliação do som adequado.

A Base Nacional Comum Curricular (BNCC) discorre sobre o ensino de pronúncia no eixo Oralidade, ao destacar que “envolve as práticas de linguagem em situações de uso oral da língua inglesa, com foco na compreensão (ou escuta) e na produção oral (ou fala) [...] **com ou sem contato face a face**¹” (BRASIL, 2018). Nestes contextos, em que as práticas orais não acontecem face a face (como em filmes, programas de TV ou da *Web*, músicas, publicidades, ligações, vídeo chat, etc) a compreensão está atrelada principalmente ao contexto e aos usos da linguagem, às temáticas e a suas estruturas (ibidem, 2018).

Assim, usar o reconhecimento de voz (RV) pode fazer parte de uma abordagem autônoma, que prioriza o aluno como maior responsável pela própria

¹ Grifo nosso.

aprendizagem. Trata-se, pois, de abordagem extremamente importante em contexto remoto, dado que o professor não consegue mediar acuradamente o grau de evolução do aprendizado do aluno, tal qual o faz no ensino presencial.

A relevância desta pesquisa está em contribuir com a comunidade docente, além de possíveis interessados no tema, apontando uma abordagem pouco explorada para o ensino de pronúncia, mormente em contextos de ensino a distância. O objetivo deste trabalho, portanto, propõe a prática da reflexão acerca do processo de ensino-aprendizagem da pronúncia de uma LE, em contexto remoto e, desta forma, oferece a professores, alunos e gestores da educação, a alternativa da aquisição fonológica da língua inglesa através do uso do reconhecimento de voz.

1. Reconhecimento de voz (RV)

O funcionamento do reconhecimento de voz se faz por meio de indexadores que previamente armazenam informações, e que podem ser essencialmente diferentes um dos outros. De forma simples, indexadores baseiam-se em áudios e transformam o arquivo de voz em outro formato de arquivo, o qual preserva as palavras e suas posições na frase.

Há duas abordagens no indexamento do áudio: indexamento baseado em texto e indexamento baseado em fonemas (SEN, DUTTA, DEYS, 2019). Os autores descrevem o indexamento baseado em texto como um grande vocabulário de reconhecimento de fala contínuo (*Large Vocabulary Continuous Speech Recognition – LVCSR*). Em resumo, trata-se de um “grande dicionário” em que palavras podem ser identificadas através de um algoritmo, e, deste modo, a voz pode ser convertida em texto. Essa abordagem tende a ser mais assertiva do que aquela baseada em fonemas, pois é capaz de prever padrões diferentes de sequências de palavras. Já os indexadores baseados em fonema funcionam com sons, ao invés de converter a fala em texto. Nessa abordagem, inicialmente, os sons de um recorte de áudio são analisados para criar um indexador com base fonêmica. Em seguida, um dicionário, contendo dezenas de fonemas, é usado para converter os termos de pesquisa do usuário em uma possível sequência de fonemas (ibidem, 2019).

De modo geral, o reconhecimento de voz pode enfrentar alguns problemas, ainda segundo Sen, Dutta e Deys (2019). Dentre essas dificuldades, destacam-se as que estão relacionadas a barulhos externos, inerentes à natureza e que podem provocar uma taxa de erro de 5 a 50%. O uso de palavras novas, que não fazem parte do “dicionário prévio”, também pode ser um empecilho – caso o RV não reconheça a palavra, devido às limitações em seu vocabulário. Como forma compensatória, ele pode sugerir palavras semelhantes, baseadas em sua base de dados, ou pode ocorrer ainda que o usuário venha a inserir esta nova palavra, e torná-la parte do dicionário. Em casos raros, o RV não encontra relação entre a palavra dita e o dicionário. Por essa razão, acaba combinando mais de uma palavra, podendo, por exemplo, trocar “*Call Paul*” por “*Calpol*”. Para além destes problemas, e considerando-se os indexadores baseados em fonemas, somente haverá assertividade nas frases mais comuns, cuja sequência dos fonemas já é previamente padronizada. Isto implica na necessidade de correção manual no resultado do reconhecimento de voz, quando em situações de alto nível de conhecimento da língua.

1.1. O reconhecimento de voz e ensino da pronúncia

Em meados de 1990, o Reconhecimento de Voz Automático (*Automatic Speech Recognition - ASR*) já havia se desenvolvido, a ponto de ser usado no aprendizado de línguas e na avaliação do inglês falado, adotado em testes orais de proficiência na Califórnia (HINCKS, 2001). Estes testes eram realizados através de uma ligação telefônica monitorada por um computador, em que o objetivo era testar o inglês de funcionários e estudantes de determinadas organizações. Em sua pesquisa, Hincks investigou os resultados de um teste completamente automatizado, chamado *PhonePass*. Na situação investigada por Hincks, este teste foi aplicado antes e depois de uma sequência de instruções realizadas com alunos, submetidos à aprendizagem do inglês através de um software, que, por sua vez, propunha um sistema de diálogos, e entregava um feedback das frases, palavras e fonemas. O *PhonePass* ditaria se esta metodologia teria sido ou não benéfica ao

desenvolvimento da língua falada dos alunos. Hincks (2001) observou que os objetivos propostos não foram alcançados, além de revelar grandes problemas com o teste. O *PhonePass* não foi eficiente na métrica do ritmo dos sons e não considerou a presença aleatória de armadilhas fonéticas, como mencionado anteriormente no exemplo de “Call Paul/Calpol”.

O feedback é possivelmente uma das maiores vantagens do uso do reconhecimento de voz, pois possibilita um retorno da frase, palavra ou texto (HINCKS, 2003; ELIMAT & ABUSEILEEK, 2014). Através deste retorno, o professor pode trabalhar aspectos que desviam do objetivo inicial, uma vez que saberá exatamente no que intervir. Contudo, a maioria dos professores ainda não teve a oportunidade de desenvolver a pronúncia adequada, de modo a usá-la como modelo de ensino (HINCKS, 2003). Desta forma, o trabalho oral com uma LE tem sido muitas vezes negligenciado.

Kim (2006) estuda o uso do software “*FluSpeak*” na instrução explícita de pronúncia. Este programa que promete trabalhar aspectos fonéticos e prosódicos da língua inglesa, e conclui que o uso do reconhecimento de voz, ainda que seja eficaz por um lado, deixa a desejar em outras esferas. A precisão do julgamento quanto à entonação pelo *FluSpeak*, quando comparado com falantes nativos do inglês, não é confiável. O *FluSpeak* parece não dar conta da variação de tons da fala, tendo em vista que leva em consideração sua base de dados e generaliza os resultados. É o que diz Hincks (2003) sobre o funcionamento de tecnologias que usam o reconhecimento de voz automático, pois trata de uma comparação probabilística entre o sinal recebido e o que se conhece sobre os fonemas armazenados na base de dados, sendo a prosódia uma das maiores limitações do RV. Contudo, tanto Kim (2006) quanto Hinks (2003) consideram que o uso do reconhecimento de voz pode ser usado efetivamente nas aulas para o desenvolvimento de habilidades orais.

Outro aplicativo, que instiga a pesquisa de Hinks (2003), é o *Talk to Me*. Este aplicativo interage com o usuário através da construção de diálogos orais. A condição mínima para que a simulação flua é que o sistema reconheça a resposta proferida dentre as 3 (três) possíveis para cada pergunta. Ou seja, há um padrão de resposta, o que facilita a busca na base de dados indexada. Se a resposta não for entendida pelo computador, o usuário precisa repeti-la até ser entendido. “Cada

diálogo consiste em trinta telas de perguntas-e-respostas, acompanhadas de ilustrações e às vezes de vídeo clipes²” (HINKS, 2003:7).

Além de frases, o software também trabalha os níveis da palavra e do fonema. Nestes exercícios, o aluno deverá repetir palavras e frases, várias vezes, até que alcance escores melhores. Os resultados da pesquisa apontaram que o reconhecimento de voz não é confiável sem a presença de um tutor humano para diagnosticar e corrigir os erros dos alunos, dado que o *Talk to Me* apontava o desvio do padrão, mas não indicava a forma adequada. Assim, poucos foram os avanços dos alunos submetidos a tal metodologia. Deste modo, entende-se que a forma de feedback deveria ser desenvolvida de maneira mais eficiente. Entretanto, os estudantes expressaram satisfação com o aplicativo, porquanto muitos afirmaram ter a possibilidade de, a partir da experiência com o software, dar continuidade aos estudos individualmente.

Hinks (2003) afirma que o uso de softwares para o aprendizado de línguas é efetivo nos níveis iniciais, quando os aprendizes normalmente carregam um sotaque forte, oriundo de suas línguas maternas. No entanto, pode ser insuficientemente efetivo para alunos de nível intermediário ou superior. Como forma de conclusão, é possível afirmar que, apesar das limitações, os programas que usam o reconhecimento de voz podem ser usados como ferramentas valiosas no ensino de pronúncia (KIM, 2006).

Silva e Silva (2017) considera o professor como fundamental nos processos de aprendizagens que fazem uso do reconhecimento de voz. De acordo com ele, “para que o uso dessas ferramentas possa auxiliar no processo cognitivo, deve-se repensar como nossos alunos se relacionam com os diversos tipos de tecnologias presentes em suas vidas” (SILVA E SILVA, 2017, p. 32), cabendo ao docente maximizar esses usos em prol da aprendizagem. O autor defende ainda que softwares assim comparam a produção do aluno com a do falante nativo e ao fornecer informações de sua produção, possibilita a melhoria do desempenho oral. “Esse recurso pode auxiliar principalmente aqueles que procuram um sotaque mais próximo do de um nativo ou que têm muitas dificuldades em serem compreendidos devido à influência da língua materna na pronúncia da segunda língua” (idem). É sob esta perspectiva que se finda o desenvolvimento dos padrões orais de uma língua

² Tradução nossa

estrangeira através do RV, e que, inclusive, em contextos onde professor e aluno se conectam somente no virtual, a autonomia do alunado precisa ser desenvolvida antes, ou mesmo em consonância, de qualquer prática de aprendizagem virtual.

2. A autonomia em tempos de pandemia e a urgência das metodologias ativas

Diante das dificuldades relativas à nova realidade remota, e devido às medidas protetivas de distanciamento social, a prática oral de LE parece por vezes impraticável. Tal percepção se destaca nas agências de ensino, as quais se mostram impossibilitadas de usar softwares, para metodologias ao vivo, decorrente de fatores que fogem da alçada da gestão. Nesse sentido, na tentativa de evitar maiores dificuldades, o professor necessita trabalhar também em um ambiente de aprendizagem no qual predomine a autonomia de aprendizagem de seus alunos, já que estes precisam aprender a desenvolver suas próprias técnicas de estudo, tendo seus potenciais explorados. Se assim ocorrer, o aluno tornar-se-á, gradativamente, mais responsável pela evolução de seu processo de aprendizado, e ainda mais ativo na construção de seu próprio saber. Hoje, mais que nunca, os modelos do professor como foco, centro e fonte única da “ensinagem” precisam ser desfeitos e desmistificados.

A autonomia dos aprendizes pode ser trabalhada através das metodologias ativas. Bastos (2006) define Metodologias Ativas como “processos interativos de conhecimento, análise, estudos, pesquisas e decisões individuais ou coletivas, com a finalidade de encontrar soluções para um problema”.

Por essa ótica, o professor passa a atuar como facilitador ou orientador do processo de ensino-aprendizagem, e, desta maneira, deverá estimular o aluno a refletir sobre seu próprio percurso de aprendizagem. Por conseguinte, os aprendizes terão mais liberdade para tomar suas próprias decisões, isto é, serão autorizados a escolher seus principais objetivos de aprendizagem, e quais as estratégias lhes são mais adequadas em sua busca pelo conhecimento ou por alguma forma de habilidade.

Disto decorre uma idealização do aprendiz, que passa a estar efetivamente engajado em seus percursos de aprendizagem; participante ativo, num processo de

interação entre pessoas, na busca da solução de problemas. O professor, por seu turno, se torna um ser um facilitador e orientador neste processo, e o foco é a “aprendizagem”. Resumidamente, o aluno não deverá mais restringir-se ao papel de um indivíduo passivo, em seu contexto de ensino-aprendizagem, apenas ouvindo e seguindo as orientações do professor. Antes, cabe ao aluno tecer suas metas, estratégias e opções para atingir o que objetivou aprender de forma mais ativa, reflexiva e autônoma.

A metodologia ativa pode ser trabalhada a distância sem muitos empecilhos, já que o objetivo é sempre o mesmo, a autonomia dos aprendizes. A quinta (5ª) competência específica de língua inglesa para o ensino fundamental da BNCC diz que se deve “utilizar novas tecnologias, com novas linguagens e modos de interação, para pesquisar, selecionar, compartilhar, posicionar-se e produzir sentidos em práticas de letramento na língua inglesa, de forma ética, crítica e responsável” (BRASIL, 2018: 244). Consideramos esta competência viável a qualquer nível de escolaridade, pois, de fato, as TDIC são facilitadoras, e, naturalmente, no ensino remoto, passam a ser inevitáveis. Assim, na nova conjuntura educacional, a prática da metodologia ativa se intensifica, e torna-se essencial para que haja processos de ensino-aprendizagem cada vez mais bem sucedidos.

Ainda sobre o eixo oralidade, a BNCC coloca que:

Para o trabalho pedagógico, cabe ressaltar que diferentes recursos midiáticos verbo-visuais (cinema, internet, televisão, entre outros) constituem insumos autênticos e significativos, imprescindíveis para a instauração de práticas de uso/interação oral em sala de aula e de exploração de campos em que tais práticas possam ser trabalhadas. Nessas práticas, que articulam aspectos diversos das linguagens para além do verbal (tais como o visual, o sonoro, o gestual e o tátil), os estudantes terão oportunidades de vivência e reflexão sobre os usos orais/oralizados da língua inglesa. (BRASIL, 2018: 241).

Com isso, percebe-se que há um forte incentivo ao uso de recursos tecnológicos para o ensino de língua inglesa. Assume-se que essa prática deverá viabilizar também maior alcance no que tange aos aspectos nativos da cultura e do idioma-alvo.

3. O aplicativo *ELSA Speak*

O aplicativo *ELSA Speak* é considerado um instrutor virtual de pronúncia de inglês. No site de venda desta plataforma, tem-se que: “Ele é capaz de ouvir a forma como os estudantes do idioma pronunciam palavras, frases ou conversas, identificar seus erros com exatidão e oferecer feedbacks preciosos em tempo real sobre pronúncias incorretas, com sugestões específicas de aprimoramento”.

O *ELSA Speak* permite ao usuário o monitoramento de seu progresso e a possibilidade de seguir um cronograma de estudos adaptado ao seu nível de proficiência em inglês, pois conta com um teste de nivelamento opcional que direciona o estudo.

O software é vencedor da competição SXSWedu® 2016 *Launch*, reconhecido como melhor *app* de aprendizado digital pela *Reimagine Education*. Ele foi considerado como um dos melhores entre os 4 (quatro) utilizados por empresas de Inteligência artificial, capaz de mudar o mundo, segundo a Forbes. Portanto, acreditamos no potencial do aplicativo em gerar mais engajamento e tornar o processo de ensino-aprendizagem mais lúdico e interativo.

Por este viés, nosso trabalho busca propor, através das TDIC, o incentivo ao trabalho com a aquisição fonológica de língua inglesa, demonstrando ser possível esta prática, mesmo em contexto de ensino remoto. Objetivamos, assim, proporcionar ao aluno a imersão neste processo mais autônomo de aprendizagem, que lhe poderá ser bastante útil, como opção para a consolidação dos aspectos relacionados à pronúncia da língua inglesa. Em outras palavras, pretendemos ofertar ao aprendiz de língua inglesa uma maneira menos tradicional, e mais dinâmica de realizar o aprendizado da pronúncia desta LE.

Nesta tentativa, propusemos uma metodologia de aprendizagem em que o aluno pudesse identificar em si mesmo seu potencial para realizar empreendimentos de aprendizagem de modo mais independente da instrução presencial do professor, e, por conseguinte, vir a ser um aprendiz mais autossuficiente na busca pela construção do próprio conhecimento. A seguinte metodologia preocupou-se em ser o mais viável possível, levando em consideração as dificuldades enfrentadas por alunos e educadores, quanto ao acesso a “tecnologia basilar”, além de propor feedback e incentivo à crescente autonomia do aluno.

4. Metodologia

Para testar os efeitos do trabalho com o aplicativo *ELSA Speak*, como alternativa para aprendizagem da pronúncia da língua inglesa, em contexto de sala de aula invertida, realizamos uma pesquisa no formato de estudo de caso.

Este trabalho contou com a participação de 02 (dois) alunos do oitavo ano do ensino fundamental de uma escola pública da rede municipal de ensino, do interior de Pernambuco. Estes alunos foram participantes voluntários da pesquisa, com a devida anuência e autorização de seus pais.

O professor-pesquisador da turma apresentou a proposta de pesquisa aos alunos durante uma das aulas remotas regulares, e em seguida recrutou dois dos estudantes, que se colocaram à disposição para colaborar com a pesquisa. Sob a orientação do professor-pesquisador, os participantes instalaram em seus *smartphones* o aplicativo *ELSA Speak* e durante um período de 10 (dez) dias o utilizaram de forma autônoma e livre.

Neste período, houve dois encontros via *Google Meet* entre o professor-pesquisador e os dois participantes, a fim de averiguar o processo e esclarecer possíveis dúvidas. O primeiro encontro ocorreu dois dias após a instalação do aplicativo, e o segundo ocorreu no nono dia de uso. Após o período de dez dias, os participantes da pesquisa responderam a um questionário no *Google Forms*, que originou os dados a serem discutidos. Seguem, abaixo, as perguntas constantes do referido questionário:

P1: Sobre a instalação do aplicativo *ELSA Speak*, houve alguma dificuldade?

() sim () não

P1.1: Se sua resposta anterior foi "sim", diga qual/quais foi/foram a(as) dificuldade(s) na instalação.

P2: Quanto ao aplicativo em si, você considera que foi fácil de manuseá-lo? Você o considera intuitivo?

() sim () não

P2.1: Se sua resposta anterior foi "não", explique por quê.

P3: Você já tinha praticado a pronúncia da língua inglesa com o suporte de algum tipo de recurso semelhante ao *ELSA Speak*?

() sim () não

P3.1: Se sua resposta foi "sim", de que maneira você exercitou sua pronúncia em inglês com recursos de aplicativos semelhantes ao *ELSA Speak*, através da internet?

P4: ANTES de fazer uso do aplicativo *ELSA Speak* para prática da pronúncia da língua inglesa, você considera que sua pronúncia estaria em que nível?

- ☐ Fraca (pouco compreensível pelo ouvinte e com muito pouca fluência de fala);
- ☐ Média (razoavelmente compreensível pelo ouvinte, porém com pouca fluência de fala);
- ☐ Boa (bem compreendida pelo ouvinte, porém com pouca fluência de fala);
- ☐ Ótima (bem compreendida pelo ouvinte e com boa fluência de fala).

P5: DEPOIS de fazer uso do aplicativo *ELSA Speak* para prática da pronúncia da língua inglesa, você considera que sua pronúncia estaria em que nível?

- ☐ Fraca (pouco compreensível pelo ouvinte e com muito pouca fluência de fala);
- ☐ Média (razoavelmente compreensível pelo ouvinte, porém com pouca fluência de fala);
- ☐ Boa (bem compreendida pelo ouvinte, porém com pouca fluência de fala);
- ☐ Ótima (bem compreendida pelo ouvinte e com boa fluência de fala).

P6: Quanto tempo, em média, você fez uso do aplicativo *ELSA Speak* para o aprimoramento de sua pronúncia?

- ☐ De quinze a trinta minutos por dia;
- ☐ Em torno de uma hora por dia;
- ☐ De duas a três horas por dia;
- ☐ Não houve organização de tempo para o uso do aplicativo diariamente.

P7: Você concorda que o aplicativo *ELSA Speak* o(a) ajudou a melhorar sua pronúncia na língua inglesa?

- ☐ sim ☐ não

P7.1: Justifique sua resposta anterior.

P8: Qual/quais foi/foram o(s) som/sons mais difícil/difíceis para você articular?

P9: Você considera que o feedback oferecido pelo aplicativo o(a) auxiliou em seu processo de aprimoramento da pronúncia de alguns sons da língua inglesa?

- ☐ sim ☐ não

P9.1.1: Justifique sua resposta anterior.

P10: Você acredita que utilizar aplicativos, com formato semelhante ao *ELSA Speak*, pode ajudá-lo(a) a aperfeiçoar sua pronúncia do inglês de forma mais autônoma e independente da instrução presencial de um(a) professora(a) de inglês?

- ☐ sim ☐ não ☐ talvez

P11: Você concorda que professores de língua inglesa deveriam motivar os alunos a utilizarem aplicativos como o *ELSA Speak* autonomamente e fazer o uso deles em aulas regulares?

P11.1: Justifique sua resposta anterior.

A partir da coleta de informações, realizou-se a interpretação dos dados, a qual será discutida a seguir.

5. Resultados e discussões

Os resultados, obtidos por meio das respostas dos alunos ao questionário, acerca da experiência com o uso do aplicativo *ELSA Speak*, buscaram entender se

o reconhecimento de voz pode, de fato, configura-se em uma ferramenta auxiliar para o professor de língua inglesa em contexto de aulas remotas.

Os questionamentos veiculados em nossa pesquisa centraram-se na investigação sobre o quão eficiente podem ser alguns aplicativos para o aprimoramento da pronúncia em língua inglesa, como forma alternativa de aprendizagem desta habilidade do aprendiz. Reiteramos ainda que um fator motivador para nossas investigações resultou da inegável realidade, concernente ao crescimento de novas práticas didático-pedagógicas. Nesse novo cenário educacional, o ensino remoto assume condição de forte relevância nos currículos das mais variadas disciplinas. Como consequência, retira o professor, enquanto figura central e indispensável na viabilização dos objetivos de ensino, e oferta ao aprendiz um papel de maior protagonismo em seus empreendimentos de aprendizagem.

Destarte, nessa nova configuração do contexto de ensino-aprendizagem, o docente se vê assumindo a reponsabilidade de conduzir turmas virtuais, por meio de interações remotas, não raro assíncronas, e, no entanto, ainda com a concepção, já datada, de que a atenção maior do processo educacional deve recair sobre sua ação docente. Em contrapartida, espera-se que os alunos se mostrem como atores mais ativos, inquietos, curiosos e construtores de seus próprios projetos de aprendizagem. Parece irrefutável, assim, afirmar que, nesta nova ordem pedagógica, nos vemos diante de realidades muito mais difíceis e desafiadoras, em particular para o professor, uma vez que deverá monitorar e intervir junto ao aluno.

Diante do exposto, passamos a analisar como se deu o uso do reconhecimento de voz durante o período experienciado, através do aplicativo ElsaSpeak, observando de que forma, mesmo num curto espaço de tempo, pode ter auxiliado no desenvolvimento da pronúncia e das capacidades autodidatas dos alunos. Para tanto, faz-se necessário avaliar os dados obtidos em nossa pesquisa e interpretá-los segundo os objetivos investigativos.

Ambos os participantes da pesquisa para a **P1** (pergunta número 1) demonstraram não revelar qualquer dificuldade na instalação do aplicativo. Deduzimos, com isso, que fatores como armazenamento interno, registro ao aplicativo ou qualquer outro possível fator adverso não interferiu no primeiro contato com o software. A esse respeito, e conforme as respostas para a **P2** (pergunta

número 2), ambas afirmativas indicam que a experiência dos usuários tenha sido intuitiva e de simples manejo, evidenciando, novamente, não terem enfrentado qualquer adversidade na utilização do aplicativo adotado para a pesquisa.

Com a **P3** (pergunta número 3), visamos detectar se os participantes já estavam familiarizados com outros aplicativos similares ao *ELSA Speak*, os quais também pudessem proporcionar a prática de pronúncia da língua inglesa, com os respectivos feedbacks. Em caso de resposta afirmativa, poderíamos inferir ser esta prática mais comum do que uma eventual falta de familiaridade com esses tipos de recursos. No entanto, as repostas foram negativas, isto é, nenhum dos participantes revelou possuir conhecimento prévio sobre esta tecnologia.

Com base nestas respostas, não nos parece insensato afirmar que muitas metodologias para o ensino da pronúncia em LE - mesmo nos contextos remotos, não fazem uso do reconhecimento de voz, para o desenvolvimento de pronúncia da língua-alvo. Sobre isso, é plausível pontuar uma ausência destes recursos para, ao menos, auxiliar ou instigar o aluno na busca sua própria aprendizagem.

Faz sentido, pois, associarmos esta interpretação ao que mencionamos anteriormente, quando das colocações da BNCC em relação à indispensabilidade dos recursos midiáticos para o uso/interação oral nas aulas (BRASIL, 2018).

Com efeito, o que se evidenciou, nos dados obtidos, foi o fato de que, mesmo em um momento histórico, em que as TDIC se mostram capazes de fornecer as necessárias conexões entre professor, aluno e sala de aula, tais recursos tecnológicos ainda carecem ser devidamente explorados.

Para as **P4** (pergunta número 4) e **P5** (pergunta número 5) interessava-nos descobrir se houve significativos avanços na pronúncia dos participantes. Como resposta à **P4**, um dos voluntários afirmou considerar sua pronúncia, antes do período de uso do aplicativo, como “fraca”. Ou seja, pouco compreensível pelo ouvinte e com muita pouca fluência de fala.

Julgamos, com isso, que nosso participante ainda esteja em nível muito básico de proficiência na habilidade da produção de fala em inglês. Por outro lado, o outro respondente do questionário afirmou considerar sua pronúncia, antes da utilização do aplicativo, como “média”, isto é, razoavelmente compreensível pelo ouvinte, porém com pouca fluência de fala.

Enquadramos, assim, este segundo participante em um nível um pouco mais avançado do que seu parceiro, porém ainda em um patamar mais básico de proficiência de fala em língua inglesa. Ao compararmos as respostas da **P5**, em que ambos consideram suas pronúncias como “média”, após o uso do aplicativo por 10 dias, percebemos que somente o participante que declarou considerar sua pronúncia como “fraca”, em um momento anterior ao uso do aplicativo, foi capaz de perceber algum grau de evolução em tal habilidade.

A esse respeito, parece-nos cabível inferir que o fator tempo de uso do aplicativo tenha interferido nesta variável. Acreditamos, portanto, que, caso nossos informantes tivessem sido submetidos a um período maior de tempo no uso do aplicativo, o resultado poderia ter sido outro. Diante destas conclusões, consideramos que os alunos, em nível mais iniciante, podem identificar resultados positivos de forma mais imediata, com o uso de recursos como os deste tipo de aplicativo. Além disso, nos foi razoável deduzir que o tempo de uso do aplicativo, nesta pesquisa, pode não ter sido suficiente para que ambos os informantes tenham realmente detectado relevante evolução em sua capacidade de pronunciar palavras na língua inglesa de forma inteligível e adequada.

Podemos ainda relacionar estes dados às respostas da **P6** (pergunta número 6), que investiga o tempo de uso diário do aplicativo *ELSA Speak*. O participante, que não percebeu evolução expressiva em sua pronúncia, fez uso do *app* numa média de quinze a trinta minutos diários. Diferentemente deste, o segundo informante, que verificou progresso em sua pronúncia, revelou não ter se planejado para dedicar tempo diário ao uso do aplicativo. Este participante ressaltou o fato de que, a depender de sua disponibilidade, teria praticado a pronúncia com auxílio do *ELSA Speak*, de forma excessiva, insuficiente ou mesmo sequer ter feito uso do aplicativo.

A **P7** (pergunta número 7) revela que os participantes concordaram quanto ao fato de que o aplicativo os teria ajudado a melhorar a pronúncia na língua-alvo, tendo um deles justificado sua resposta ao dizer: “*Antes tinha muita dificuldade com algumas palavras, esse app me ajudou a ‘compreender’ melhor...*”

Nossa interpretação da justificativa deste informante foi que as dificuldades mencionadas poderiam estar relacionadas tanto à pronúncia, como à percepção das referentes palavras produzidas por terceiros. Isto se fez mais claro quando este

informante justifica: “...pois ele ajuda a falar as palavras da maneira certa, e incentiva no aprendizado”. Percebemos, assim, que este voluntário talvez considere que o aplicativo realmente o tenha incentivado na própria busca pelo aprimoramento da pronúncia em inglês. Vale lembrar que esta segunda resposta foi dada pela pessoa que percebeu evolução na própria pronúncia. Porém, cabe lembrar que este também foi o participante que afirmou não ter organizado previamente seu tempo de uso do aplicativo.

Na **P8** (pergunta número 8), objetivamos identificar quais dos sons, praticados com o auxílio do *ELSA Speak*, os participantes consideraram mais difíceis de articular. Para um deles, foi o que contava com a presença gráfica de “oo”, cujo som normalmente equivale ao fonema /ʊ/ (como em “good” /gʊd/). Já o segundo informante indicou terem sido os sons dos correlatos ortográficos “th” e do “ph” aqueles que exigiram mais esforço, já que o aluno os teria considerado como mais difíceis de articular, particularmente quando pronunciados em frases mais extensas. Esta informação também foi constatada durante a primeira reunião do projeto. Naquele momento, este participante já mencionou o fato de que a produção do “th” lhe seria especialmente mais complicada. O professor-pesquisador precisou, assim, intervir e explicar melhor e mais detalhadamente a forma como devemos articular este fonema na língua inglesa, tendo sido necessário mostrar visualmente o correto ponto e modo de articulação deste som, para a produção adequada do referido fonema, e para a devida distinção do / ð / e / θ /.

Podemos concluir, com isso, que o *ELSA Speak* ainda requer algum aperfeiçoamento, no sentido de complementar sua efetividade no suporte ao aprendizado da pronúncia. Isto porque nos parece pertinente sugerir que, uma vez inserido algum recurso tecnológico, capaz de exemplificar visualmente as produções sonoras dos fonemas (quando e se solicitado pelo usuário), tal adequação técnica provavelmente redundaria em resultados ainda exitosos nas produções fonéticas dos aprendizes.

Ainda assim, embora tenham apresentado certas dificuldades na produção de alguns sons, ambos os informantes afirmaram entender como de grande importância a oferta de feedback por parte do aplicativo. Como observado na resposta da **P9** (pergunta número 9), nossos dois participantes concordaram ao destacar o fato de que este retorno os auxiliou positivamente, na prática da pronúncia dos sons da

língua-alvo. Tal percepção foi verbalizada em palavras, nas seguintes justificativas: *“me ajudou bastante”* e *“me ajudou a falar as palavras com a pronúncia correta”*.

Hincks (2003) e Elimat & Abuseileek (2014) reconhecem o feedback como a maior vantagem em se utilizar o RV. No aplicativo em questão, apenas os alunos participantes tiveram acesso a esse feedback. Deduzimos, pois, que estes alunos possivelmente não souberam como tirar proveito deste recurso.

Ao responde à **P10** (pergunta número 10), que buscou verificar as impressões das experiências e práticas de pronúncia com o *ELSA Speak*, um dos informantes nos revelou acreditar que aplicativos, com formato semelhante ao *ELSA Speak*, podem ajudá-lo a aperfeiçoar sua pronúncia em inglês de forma mais autônoma e independente da instrução presencial de um professor.

Esta resposta coincide com a postura do aluno voluntário, que percebeu uma evidente evolução em sua pronúncia, após o período de uso do aplicativo. Por outro lado, o participante que persistiu em considerar sua pronúncia como “média”, marcou a opção “talvez”, para essa pergunta. Entendemos, assim, ser viável inferir que nosso aplicativo tem a potencialidade de oferecer resultados benéficos aos usuários, independentemente do acompanhamento de instruções presenciais de um professor. Contudo, com esta resposta (“talvez”), também nos compele a deduzir que o aluno pode, eventualmente, demonstrar insegurança, ao fazer uso apenas do aplicativo para o aprendizado da pronúncia do inglês, de forma autossuficiente e sem o amparo e os esclarecimentos de um professor.

Entretanto, e de forma contraditória a esta interpretação, os dois respondentes à **P11** (pergunta número 11) estavam de acordo, ao dizer que os professores de língua inglesa deveriam motivar os alunos a utilizarem aplicativos como o *ELSA Speak* autonomamente e segundo a disponibilidade de cada um. Contudo, ambos reiteraram entender que o professor também deveria fazer o uso desses recursos tecnológicos em suas aulas regulares.

Como justificativa para essa pergunta, os informantes destacaram em suas falas o seguinte: *“ajuda muito no inglês”* e *“muitos acabam tendo bastante vergonha para perguntar ou até mesmo ‘estudar’ com um professor, então acredito que se usássemos esse app mais vezes... seria bem melhor!”*.

Esta segunda justificativa adequa-se ao que diz Hincks (2003), quando este defende que muitos alunos se sentem desconfortáveis ao se perceberem como o centro das atenções nas salas de aula.

Por essa perspectiva, cabe ressaltar a importância e os benefícios oriundos do uso regular e devidamente monitorado de aplicativos de reconhecimento de voz. Isto porque tais recursos da TDIC podem se configurar como excelentes alternativas para todos os aprendizes de uma LE, em especial, para aqueles que se mostram frequentemente mais inseguros em se comunicar na língua-alvo. Estes estudantes, com frequência, revelam-se inconscientes de suas próprias potencialidades na busca pelo progresso em seu nível de proficiência na LE em estudos.

Assim, fato é que estes aplicativos têm mostrado, ao longo dos últimos anos, significativos benefícios para os alunos que desejam dar continuidade a seus estudos, mas que preferem fazê-lo por conta própria, e contando com suas melhores estratégias de aprendizagem. Com isso, estudantes de uma LE, com semelhante perfil, podem tirar melhor proveito de seu tempo de aprender, e obter, assim, resultados mais sólidos de seus esforços, com o suporte desses aplicativos inovadores para o aprendizado de línguas, sobretudo quando se trata de aprendizado dos padrões fonético-fonológicos da língua inglesa.

6. Conclusão

Diante dos resultados obtidos, embora numa amostra reduzida e num período curto de uso, concluímos fundamentalmente que o aplicativo *ELSA Speak*, quando utilizado como meio de aprimoramento da pronúncia da língua inglesa, sem a assistência, o acompanhamento e as orientações do professor, não demonstrou grande impacto sobre os usuários, participantes desta pesquisa. Contudo, é necessário ressaltar que, para um deles, o uso do aplicativo com tal finalidade tenha se revelado mais eficiente do que para o outro. Parece-nos, assim, bastante razoável presumir que o tempo de uso do *ELSA Speak*, pelos dois alunos de nossa pesquisa, não tenha sido suficiente para que ambos atingissem maior nível de evolução em seus processos de aprendizagem. Ademais, importa também salientar nossa crença de que a participação efetiva e contínua do professor, junto aos

aprendizes, em projetos de ensino que também incluam softwares para aprendizagem de uma LE, ainda é de grande relevância e inegável influência. Resumidamente, acreditamos ser necessário haver maior intervenção do professor durante o todo e qualquer empreendimento de aprendizagem; seja de forma virtual ou híbrida.

Nosso projeto visou primordialmente investigar as vantagens e os benefícios proporcionados aos alunos de LE, que se valem dos recursos tecnológicos das TDIC, para o aperfeiçoamento de sua habilidade de pronúncia na língua-alvo.

A despeito deste fato, os dados que nos foram fornecidos, nos autorizam concluir que, sem a devida supervisão e orientação do docente, este propósito tende a não ser alcançado satisfatória e plenamente.

O software utilizado demonstrou ser mais eficiente para os alunos em níveis mais básicos de aprendizagem. Por essa constatação, entendemos que o mesmo deverá ocorrer para grande parte de aprendizes, em níveis semelhantes de proficiência. Ante a esta conclusão, ratificamos novamente a necessidade de que o professor se faça mais presente, para oferecer ao aprendiz orientação e feedback, ao longo da implementação de tais práticas de pronúncia, em que se inserem aplicativos para o auxílio na aprendizagem.

No caso do ELSA Speak, estes dados não são compartilhados com o professor. Não obstante, é possível pensar em alternativas para que o aluno desenvolva suas próprias estratégias, no sentido de tentar alcançar feedbacks mais produtivos, elucidativos e capazes de instrumentalizá-lo, a fim de que possa dar continuidade a seus estudos, sem grande dependência da figura do professor.

Contudo, não podemos negligenciar o enorme potencial para a prática da pronúncia em uma LE, com o uso da inteligência artificial do *ELSA Speak*. Ainda assim, não nos permitimos afirmar que apenas e tão somente com o uso do referido aplicativo, especialmente se for utilizado por pouco tempo, haja de fato notáveis diferenças e progressos na habilidade da pronúncia.

Cabe reforçar, a esse respeito, que não acreditamos na eficiência do aplicativo, se este não estiver associado a uma metodologia de ensino bem embasada e planejada, como forma de apoio aos usuários, em aulas regulares de língua inglesa.

Inegável o fato de que existem grandes vantagens na utilização desses tipos de aplicativo de aprendizagem. Além da praticidade, agilidade, rápida interação e forte apelo afetivo (em especial para os jovens aprendizes), ainda se nos parece bastante evidente e revelador a indispensável presença de um professor durante o projeto todo e qualquer projeto de ensino. Idealmente, este professor deveria ser capaz de atuar, não apenas como instrutor ou orientador do aluno, mas, antes de tudo, como agente motivador ao longo de todo o processo educacional. Ou seja, um professor que se mostre, sobretudo, como um guia para seus alunos; alguém com a habilidade de fazer com que seus aprendizes possam identificar seus próprios recursos e estratégias de aprendizagem. E, desta forma, deveremos adentrar em um novo contexto de ensino-aprendizagem de línguas.

Nessa nova orientação pedagógica do terceiro milênio, entendemos que o trabalho com recursos tecnológicos, ofertados por aplicativos inovadores de reconhecimento de voz, tais como o do *ELSA Speak*, ainda não pode ser visto como um tipo de solução mágica para as aprendizagens remotas. Antes, porém, estes novos recursos da tecnologia, já há bastante tempo, provaram-se credenciados a atuar como aliados da educação a distância, na tentativa de reduzir barreiras para o bom êxito no aprendizado.

Nerbonne (2003) assume que o RV não é perfeito. O autor vai além, e afirma que isso provavelmente não será uma realidade num futuro imediato. Todavia, ainda para este autor, o RV deve e merece ser visto com destaque e seriedade. Isso se, de fato, os gestores educacionais pretendem um ganho para todos os atores envolvidos nos investimentos em educação, e, no que tange ao escopo de nossa pesquisa, no desenvolvimento de metodologias que vislumbrem utilizá-lo de forma sistemática, na aprendizagem de línguas estrangeiras.

Impinge-nos, pois, aceitar, como sociedade, o fato de que tal cenário já está se fazendo presente na realidade educacional de nosso país, mesmo que ainda de forma coadjuvante. Em suma, não é irreal antever um contexto educacional institucionalizado, no qual haja o uso regular, frequente e complementar dos recursos tecnológicos para a aprendizagem. Em assim sendo, deveremos formar uma geração de novos aprendizes, gradativamente mais independentes, autônomos, ativamente engajados em seus objetivos de aprendizagem, e cientes de

que é deles o papel central e a responsabilidade maior, na busca por resultados mais satisfatórios para seus empreendimentos de aprendizagem.

REFERÊNCIAS

BASTOS, C. C. **Metodologias ativas**. 2006. Disponível em: <http://educacaoemedicina.blogspot.com/2006/02/metodologias-ativas.html> Acesso em: 27/10/20.

BRASIL. Base Nacional Comum Curricular: Ensino Médio. Brasília: MEC/Secretaria de Educação Básica, 2018. Disponível em: <http://basenacionalcomum.mec.gov.br/>. Acesso em: 22/08/2020.

CANI, J. B., E. G. C. SANDRINI, G. M. SOARES, e K. SCALZER. Educação e covid-19: a arte de reinventar a escola mediando a aprendizagem 'prioritariamente' pelas TDIC. **Revista Ifes Ciência**, vol. 6, nº 1, junho de 2020.

ELIMAT, A. K.; ABUSEILEEK, A. F. Automatic speech recognition technology as an effective means for teaching pronunciation. **JALT CALL Journal**, v. 10, n. 1, p. 21-47, 2014.

ELSASPEAK. Todos podem falar inglês. 2020. Disponível em <https://elsaspeak.com/br/promo/aff> . Acesso em 20/04/2021.

FREIRE, P. **Pedagogia da autonomia**: saberes necessários à prática educativa. 37. ed. São Paulo: Paz e Terra, 2008.

HANSEN, F. **O Impacto das novas tecnologias de informação e comunicação: por uma resignificação da relação professor/estudante**. In: BENEVENUTO JR.; STEFFEN, CÉSAR (orgs.). Tecnologia, pra quê? os impactos dos dispositivos tecnológicos no campo da comunicação. Porto Alegre, Armazém Digital, 2012, p. 117 -133.

HINCKS, R. Using speech recognition to evaluate skills in spoken English. **Working papers/Lund University, Department of Linguistics and Phonetics**, v. 49, p. 58–61-58–61, 2001.

_____. Speech technologies for pronunciation feedback and evaluation. **ReCALL**, v. 15, n. 1, p. 3-20, 2003.

HONORATO, H. G., MARCELINO, A. C. K. B. **A arte de ensinar e a pandemia COVID-19: a visão dos professores**. *REDE-Revista Diálogos em Educação* ISSN: 2675-5742 1.1 (2020): 208-220.

KIM, I. Automatic speech recognition: Reliability and pedagogical implications for teaching pronunciation. **Educational Technology & Society**, v. 9, n. 1, p. 322-334, 2006.

MENA, M. R. S. **A resistência do professor quanto ao uso das tecnologias e cultura digital na escola médio engenheiro Annes Gualberto**. TCC (especialização) - Universidade Federal de Santa Catarina. Centro de Ciências da Educação. Departamento de Metodologia de Ensino. Educação na Cultura Digital. 2016.

PACHECO, M. L. S.; LOPES, R. P. TIC E TRABALHO DOCENTE: DAS DIFICULDADES À RESISTÊNCIA AO USO DE TECNOLOGIAS NO PROCESSO EDUCATIVO. **Anais da Semana de Licenciatura**, Jataí, GO, p. 353-367, out. 2019. ISSN 2179-6076. Disponível em: <http://revistas.ifg.edu.br/semlic/article/view/706>. Acesso em: 16 jul. 2021.

SEN, S.; DUTTA, A.; DEY, N. **Audio Processing and Speech Recognition: Concepts, Techniques and Research Overviews**. Springer, 2019.

SILVA E SILVA, V. **Ensino Superior 2.0: O English Central como ferramenta de apoio para o ensino-aprendizagem de inglês**. Dissertação de Mestrado. Pontifícia Universidade Católica de São Paulo – Tecnologias de Inteligência e Design Digital. São Paulo, 2017.