

**FATORES A SEREM LEVADOS EM CONSIDERAÇÃO PARA O
DESENVOLVIMENTO DE TESTES DE PROFICIÊNCIA ORAL EM CONTEXTO
VIRTUAL**

**ASPECTS TO BE CONSIDERED IN THE DEVELOPMENT OF ORAL
PROFICIENCY TESTS IN VIRTUAL CONTEXTS**

ANA CRISTINA BIONDO SALOMÃO*

RESUMO: Novos contextos de aprendizagem virtual de línguas estrangeiras vêm surgindo com o desenvolvimento das novas tecnologias. Há a necessidade agora de se pensar em como a competência linguística poderá ser também avaliada em meio virtual. Este artigo traz uma retrospectiva da origem e desenvolvimento das provas de proficiência oral para falantes de línguas estrangeiras, buscando também levantar e discutir os aspectos a serem levados em consideração para a formulação deste tipo de teste em ambiente virtual.

Palavras-chave: avaliação; proficiência oral; tecnologia.

ABSTRACT: New virtual contexts for learning foreign languages have been arising from the development of new communication technologies. Now, there is the necessity to think about how the students' communicative competence may be evaluated in such virtual contexts. This article aims at presenting a retrospective of the origin and development of oral proficiency tests for speakers of foreign languages, as well as pointing out and discussing the aspects that should be taken into account when designing this kind of tests for virtual environments.

Keywords: language assessment; oral proficiency; technology.

INTRODUÇÃO

O uso de novas tecnologias, sobretudo o do computador e da internet, no ensino de línguas, tem causado grande impacto e questionamentos sobre o conhecimento e a prática dos profissionais que trabalham na área de Linguística Aplicada, professores e pesquisadores, assim como dos alunos, que agora dispõem de outros recursos para aprender línguas estrangeiras.

Muitas escolas e professores têm adotado o uso de computadores em sala de aula como atividade extracurricular para o ensino de línguas estrangeiras, e no que concerne à avaliação, seja em provas de nivelamento (*placement test*) ou provas de proficiência (como o TOEFL, por exemplo), o uso de computadores tem se tornado cada vez mais presente.

* UNESP/IBILCE, São José do Rio Preto (SP), Brasil. <anacbsalomao@yahoo.com.br>

Chapelle e Douglas (2006) afirmam que sempre que a instrução for oferecida on-line, faz sentido que os professores também pensem na possibilidade de exames (testes) on-line. Segundo os autores, há perspectivas positivas sobre as consequências do desenvolvimento das práticas de testes mediados por computador, que podem beneficiar os candidatos uma vez que cada vez mais as novas tecnologias da informação e comunicação (TIC) têm adentrado as salas de aulas.

Na área de avaliação (*language testing*) em língua estrangeira, novas questões também começam a surgir por conta das novas possibilidades oferecidas pelo uso do computador para a elaboração e execução de testes, denominado CALT (em inglês, *computer-assisted language testing*), e que denominaremos aqui Testes de Línguas mediados por computador (TLMC).

... TLMC têm o potencial para ser um tipo de efeito retroativo positivo se considerarmos os benefícios que muitos acreditam estarem associados ao acesso regular à tecnologia no aprendizado de segunda língua. Muitos professores e pesquisadores afirmam que os alunos deveriam ter acesso a materiais de ensino on-line na esperança de que tal prática aumente o letramento digital e o letramento na língua alvo (CHAPELLE; DOUGLAS, 2006, p.18).

Alguns pesquisadores vêm estudando a implementação e os efeitos da avaliação mediada por computador há mais de 20 anos. Segundo Chapelle e Douglas (2006), no *Language Testing Research Colloquium* (LTRC) de 1985, no qual aparecia somente uma minoria de pesquisadores trabalhando com avaliação em contexto tecnológico, já se previam muitas das questões que se tornariam objetos de pesquisa na área de avaliação na década seguinte, principalmente aquelas ligadas a construtos, validade e autenticidade.

Bachman e Alderson, em seu prefácio ao livro *Assessing Language through Computer Technology* de Chapelle e Douglas (2006) afirmam que a mais surpreendente mudança trazida pelas novas TIC para a área de avaliação é o potencial para a aplicação de uma variedade de testes on-line a qualquer hora, em qualquer lugar do mundo e com *feedback* imediato. Essa possibilidade gera também perguntas sobre potenciais problemas que poderão ocorrer, assim como preocupações com questões como seleção, desenvolvimento e uso de tais testes.

Chapelle e Douglas (2006) levantam alguns questionamentos relacionados à confiabilidade, validade, praticidade e autenticidade de testes realizados por meio de computador. Segundo os autores, há vários fatores que devem ser considerados juntamente com as mudanças trazidas pelas novas tecnologias, inclusive os relacionados a circunstâncias físicas e temporais, assim como de segurança. Eles comentam, ainda, trabalhos de outros pesquisadores na área de avaliação em relação às habilidades de escrita e leitura, como a discussão feita por Weigle (2002, apud CHAPELLE; DOUGLAS, 2006) sobre a avaliação da habilidade escrita mediada por computador, os meios desenvolvidos para correções e atribuições de notas por computador e os efeitos da tecnologia sobre a escrita; e o trabalho de Alderson (2000, apud CHAPELLE; DOUGLAS, 2006) sobre o desenvolvimento de métodos e o construto da leitura on-line. Entretanto, afirmam que questões relacionadas às habilidades de produção e compreensão oral por meio do computador podem ficar subanalisadas nos contextos educacionais. Os autores afirmam:

O uso da linguagem é afetado pelo contexto no qual a interação ocorre. (...) Portanto, é importante entender como os computadores podem ser diferentes de outros meios ao apresentar as tarefas de um teste e, assim também, como a tecnologia pode afetar as interpretações sobre desempenho no teste (CHAPELLE; DOUGLAS, 2006, p. 20-21).

Tais idéias e questionamentos nos levaram a pensar um pouco mais sobre a avaliação de proficiência oral em contexto virtual e a escrever este artigo. Perguntamo-nos como a tecnologia está mudando as possibilidades de testes de proficiência oral em língua estrangeira, e decidimos buscar na literatura da área de avaliação fontes que nos ajudassem a discutir a questão. Nossa reflexão foi guiada pela seguinte pergunta e suas subdivisões: 1) Que fatores devem ser levados em consideração ao desenvolver-se um teste de proficiência oral em contexto virtual, mediado por computador? a) Que aspectos devem ser adequados para este meio? b) Quais as possíveis vantagens e limitações de um teste de proficiência oral em tal contexto? c) O uso de computador neste caso pode ser inovador e ao mesmo tempo manter a validade que um teste de proficiência deve ter?

O artigo está organizado da seguinte maneira: primeiramente, resumimos a história dos testes de proficiência oral, desde sua criação, focalizando, então, a entrevista como instrumento de avaliação. Em seguida, falamos sobre as novas tecnologias da informação e comunicação e de seu impacto para o ensino e a avaliação em língua estrangeira, tratando mais especificamente, então, dos testes de proficiência oral. Ao final deste levantamento, voltamos às perguntas iniciais para discussão das mesmas segundo nossas reflexões a respeito de tal temática.

1. A HISTÓRIA DOS TESTES DE PROFICIÊNCIA ORAL

Sabe-se que a área da avaliação em segunda língua é um dos campos mais jovens da Linguística Aplicada, e a teoria e prática de testes de proficiência oral representam a subárea mais jovem no campo da avaliação.

Testes orais sempre foram considerados importantes, apesar de sua falta de confiabilidade e impraticabilidade de uso em larga escala. O crescimento do uso de tais testes emergiu da expansão educacional na década de 1920 nos EUA, de acordo com Fulcher (2003). O autor relata que inicialmente o desenvolvimento de testes de habilidade oral se deveu a razões de controle de imigração e para uso de funcionários militares. Mas foi na época da Segunda Guerra Mundial que eles se tornaram foco de interesse. O autor afirma:

Depois da Segunda Guerra Mundial foi a contínua pressão feita pela administração americana para que se registrasse as capacidades de diplomatas e militares, e preparasse o país para conflito futuro, que levaram ao programa de desenvolvimento do primeiro teste oral de larga escala no *Foreign Service Institute* (FULCHER, 2003, p. 17).

Em 1942, foi criado o “The Army Specialized Training Program (ASTP)”, um curso de línguas especialmente desenvolvido para o pessoal das forças armadas americanas com foco na habilidade oral nos campos de engenharia e medicina. Em 1944, devido à necessidade de se aplicar um teste oral em um programa como este que estava sendo desenvolvido pelo

Queen's College, em Nova York, Agard e Dunkel (1948, apud FULCHER, 2003), criaram o seu próprio teste para avaliar os alunos do ASTP. Este teste foi um dos precursores das entrevistas de proficiência oral (*Oral Proficiency Interview* – OPI).

Após a guerra, o *Foreign Service Institute* (FSI), preocupado com as demandas linguísticas necessárias para oficiais do governo americano na ocupação de postos militares em terras estrangeiras, foi estabelecido como órgão oficial para o ensino de línguas estrangeiras. Em 1952, devido à demanda de se medir a familiaridade dos oficiais com línguas e culturas estrangeiras, foi criado um comitê com a responsabilidade de desenvolver um teste de habilidade oral, no qual foram desenvolvidas as primeiras públicas e largamente usadas faixas de proficiência, que consistiam em 6 faixas (FULCHER, 2003). As faixas eram holísticas, com descritores intuitivos. Entretanto, em 1958, uma relação de cinco fatores foi acrescentada a essas faixas para auxiliar os examinadores: sotaque, compreensão, fluência, gramática e vocabulário (ADAMS, 1980, apud FULCHER, 2003).

Em 1970, os testes orais foram adotados por muitas universidades com o intuito de dar certificação a professores bilíngues, e devido ao crescente interesse na recém surgida abordagem nocional-funcional para o ensino de língua inglesa, um teste de habilidade oral foi considerado a forma de avaliação mais condizente com os novos métodos de ensino.

Ao ser usado em contextos acadêmicos, verificou-se a necessidade de adaptação das faixas de proficiência do FSI, já então standardizadas e chamadas “Interagency Language Roundtable – ILR”, para os novos propósitos de um teste de habilidade oral para ser usado fora do governo. O *American Council on the Teaching of Foreign Languages* (ACTFL) recebeu a incumbência de produzir critérios nacionais para esse fim e em 1982 lançou o *ACTFL Provisional Proficiency Guidelines*. Em 1986, foi lançada a versão completa das *Guidelines* e em 1999 uma versão revisada. Foram essas *Guidelines* que geraram o interesse e a ampla discussão sobre proficiência oral em língua inglesa, que se estende até os dias de hoje.

1.1. A entrevista como instrumento de avaliação de proficiência oral

No começo do século 20, quando se falava em avaliação da habilidade oral, segundo Fulcher (2003), estava-se normalmente referindo a questões de pronúncia e acuidade gramatical, geralmente evidenciadas por tarefas de leitura em voz alta.

Hoje em dia, os testes de proficiência oral utilizam-se largamente do formato entrevista, conhecido na literatura internacional como *Oral Proficiency Interview* (OPI), para a avaliação do desempenho na habilidade oral do candidato.

De acordo com McNamara (2000), esse tipo de entrevista consiste em uma curta interação entre o candidato e um examinador¹, na qual o seu desempenho é julgado de acordo com faixas de descritores de desempenho com diferentes níveis. A OPI teve origem na necessidade do governo norte-americano de preparar e avaliar pessoas para lidar com a comunicação oral em línguas consideradas estratégicas durante a Guerra Fria.

¹ É importante ressaltar que o texto de McNamara traz o examinador como falante nativo da língua alvo; entretanto, corroboramos as ideias de Consolo (2004) de que no contexto brasileiro não somente o falante nativo atua como entrevistador, mas também falantes não-nativos competentes na língua-alvo.

Segundo Fulcher (2003), o FSI foi um dos órgãos a conduzir este tipo de entrevista, que geralmente tinha duração de 10 a 40 minutos e era conduzida por um falante nativo. O procedimento seguido pelo entrevistador consistia em iniciar uma conversa com uma simples formulação social e a partir daí proceder à elicitación de estruturas gramaticais conhecidas e utilizadas pelo candidato. Depois disso, um jogo de papéis (*role play*) ou um diálogo preparado era usado. As técnicas de elicitación eram preparadas anteriormente. O examinador tomava notas e classificava os candidatos segundo uma escala de seis pontos em cinco fatores (sotaque, gramática, vocabulário, fluência e compreensão).

Hoje em dia, as entrevistas feitas em testes de proficiência oral em língua estrangeira adotam estilos menos estruturados e mais conversacionais. Segundo Brown (2003):

Entrevistas conversacionais são caracterizadas por formatos sem *script* e relativamente não-estruturados. Enquanto os entrevistadores geralmente recebem guias que sugerem tópicos e focos para questionamentos, questões específicas não são nem pré-formuladas nem idênticas para cada candidato; a interação deve ocorrer de maneira conversacional (BROWN, 2003, p.1).

O uso de entrevistas para a avaliação de proficiência oral tem recebido críticas, como a de Byrnes (1987, apud CONSOLO, 2004, p.272), de que elas ilustram apenas um “domínio específico de interação verbal”, e Bachman (1988), de que sua imprevisibilidade pode comprometer a confiabilidade do teste.

Por outro lado, há autores que defendem o uso de entrevistas para a avaliação da proficiência oral em língua estrangeira, como Molder e Halleck (1998, apud CONSOLO, 2004, p.272) afirmando “que as entrevistas constituem eventos de fala autênticos e que ilustram comportamentos de comunicação verbal, mesmo que esses comportamentos não sejam exatamente iguais àqueles que ocorrem em conversas mais informais”.

Apesar das críticas, a entrevista continua a ser largamente utilizada em testes de proficiência internacionalmente reconhecidos, tais como ACTFL-OPI (*American Council on the Teaching of Foreign Languages*), que pode ser feita até mesmo por telefone com hora marcada, UCLES (*The University of Cambridge Local Examinations Syndicate*), IELTS (*International English Language Testing System*).

2. NOVAS TIC, ENSINO DE LÍNGUAS E AVALIAÇÃO

O impacto das novas tecnologias da informação e comunicação (TIC) na vida particular e profissional das pessoas e os novos contextos de comunicação gerados pelas mudanças advindas delas certamente deverão nos levar a repensar o vínculo entre educação e sociedade, assim como alterar profundamente, de agora em diante, a formação das pessoas e as relações de ensino e aprendizagem, apresentando assim também reflexos sobre as formas de avaliação. Para Warschauer (2004, p. 15), “no âmbito do ensino e aprendizagem, existem obviamente muitos tipos de interação de salas de aula (ou distantes) que são possibilitadas pelos computadores e a Internet que simplesmente não poderiam ter ocorrido previamente”.

Warschauer (2000) apresenta as habilidades de análise crítica, avaliação, experimentação, colaboração, comunicação, abstração, pensamento sistemático e persuasão como fundamentais para lidar com a informação e a tecnologia e obter sucesso no mundo do século 21. Brunner (2004), citando Gardner (1999), assevera a necessidade do desenvolvimento de funções cognitivas superiores como resolução de problemas, busca pelo problema, planejamento, reflexão, criatividade e entendimento mais profundo, como preparação do profissional para o mercado de trabalho desse novo século. O autor afirma que:

...a própria experiência de uma sociedade crescentemente saturada de informação, assim com as exigências de uma individualização realizada em contextos de opções crescentes obrigam a uma maior reflexividade, não só das pessoas em seu fazer cotidiano, mas da sociedade em seu conjunto. Os indivíduos devem processar cada vez mais informação, filtrar, discriminar, pensar, valorar, reter, combinar e esquecer (BRUNNER, 2004, p.36).

Essa nova realidade trazida pelos avanços tecnológicos traz implicações para a área de avaliação em língua estrangeira uma vez que mexe com nossas concepções de comunicação e domínio de uso da linguagem, trazendo efeitos sobre o construto e o desenvolvimento de testes mediados por computador.

2.1. O uso de computadores, o ensino de língua estrangeira e a avaliação

O uso de computadores no ensino de línguas estrangeiras não é algo novo. Segundo Warschauer e Healey (1998), computadores têm sido usados no ensino de línguas desde 1960. Entretanto, formas variadas de utilização desta ferramenta foram feitas desde então, de acordo com abordagem de ensino dominante, a disponibilidade e o nível de tecnologia existentes em cada época. Desta forma, os autores dividem a aprendizagem de línguas mediada por computador (*CALL – Computer-assisted Language Learning*) em três estágios.

O primeiro estágio é denominado CALL behaviorista, situada entre as décadas de 1960 e 1970, na qual eram utilizados programas inspirados na instrução programada, com exercícios de repetição, usualmente chamados de “*drill-and-practice*”. O computador era visto como um tutor mecânico, que nunca se cansava ou julgava os alunos e permitia-lhes trabalhar em seu próprio ritmo.

No final da década de 1970 e início da década de 1980, começam a surgir propostas de uso comunicativo dos computadores, largamente influenciadas pelo surgimento da Abordagem Comunicativa e pelas teorias cognitivas que rejeitavam a visão behaviorista, defendendo a aprendizagem como um processo de descoberta, expressão e desenvolvimento. Os programas utilizados nessa segunda fase incluíam exercícios de reconstrução de texto e simulações, que visavam proporcionar a interação entre os aprendizes após o uso do computador. O foco, segundo Warschauer e Healey (1998), não estava no que os alunos faziam com a máquina, mas sim na interação que faziam longe dela enquanto trabalhavam ou após o uso do computador.

Uma nova perspectiva começou a surgir a partir da década de 1990, com a mudança de uma visão cognitiva do ensino comunicativo para uma visão mais social, sociocognitiva, que dava mais ênfase ao uso da língua em contextos sociais autênticos. Essa nova

perspectiva, advinda das abordagens de ensino baseado em tarefas (*task-based*), baseado em projetos (*project-based*) e baseado em conteúdos (*content-based*), foi chamada de CALL integrativo (WARSCHAUER; HEALEY, 1998) e procurava integrar as várias habilidades e a tecnologia no processo de aprendizagem de maneira mais completa, proporcionando aos aprendizes a interação com a máquina com uma variedade de ferramentas em processo contínuo de aprendizado e uso da língua-alvo.

O desenvolvimento e uso de testes de língua estrangeira seguiu também caminho parecido, conforme o modelo de aquisição de línguas ou abordagem de ensino dominante em cada época. Assim, Spolsky (1975, apud MARTINS, 2005) divide a trajetória de ensino/aprendizagem/avaliação em três períodos: *o pré-científico, o psicométrico-estruturalista e o sociolinguístico-integrativo*.

O período *pré-científico* se inicia a partir do século XVIII e se estende até o início da década de 1950, “caracterizando-se pelo foco na forma, na tradução e nos exames orais, com seus resultados determinados de forma impressionista” (MARTINS, 2005).

O período seguinte, referente às décadas de 1950 e 60, *psicométrico-estruturalista*, reflete a contribuição da psicologia behaviorista e da linguística estruturalista, compondo-se de “elementos isolados que fragmentavam a linguagem em categorias fonológicas, lexicais e sintáticas, mensurando o conhecimento de tais categorias sem levar em conta as correlações entre essas categorias no discurso” (MARTINS, 2005), evidenciando a preocupação com a forma, comum na época. Esse período se caracterizou por objetividade na correção e a possibilidade de se quantificar os dados com facilidade.

No período seguinte, que se iniciou juntamente com o chamado movimento comunicativo, a avaliação entrou numa nova fase, denominada de *sociolinguístico integrativo*, no qual a linguagem passou a ser vista como um código inserido em determinado contexto, usado em situações reais de comunicação. Segundo Martins (2005):

Nesse período, a avaliação vê o conhecimento de uma língua como algo que deve ser avaliado em termos da habilidade do usuário de uma língua em lidar com funções linguísticas, bem como usar a língua de maneira específica em situações definidas, visão esta que privilegiaria o desempenho do sujeito que faz uso da língua. A avaliação da linguagem oral, dentro dessa abordagem, possibilita o ensino integrativo (e não de elementos isolados), evitando a compartimentação da linguagem. (MARTINS, 2005).

Segundo Alderson, Clapham e Wall (1998), as mudanças provenientes desta nova concepção se fizeram presentes no conteúdo e no método dos testes, com o surgimento de provas baseadas em tarefas e maior ênfase na avaliação do desempenho do candidato. Além disso, os testes de múltipla escolha foram substituídos por teste de cloze e com respostas curtas e outros procedimentos de respostas mais abertas, cedendo espaço para correções mais subjetivas.

Em termos de avaliação mediada por computador, o que se viu mais frequentemente como advento do uso da tecnologia para fins de avaliação foram testes de múltipla escolha, de ordem estruturalista, oferecidos em CD-ROM, e o desenvolvimento de CAT – *computer adaptive tests*, ou seja, testes com flexibilidade e adaptabilidade, nos quais os computadores podem adaptar o insumo subsequente de acordo com as respostas anteriores dos candidatos, possibilitando a criação de um teste sob medida para o candidato (ALDERSON; CLAPHAM; WALL, 1998; CHAPPELLE; DOUGLAS, 2006).

Nos últimos anos, um modelo integrativo passou também a ser pesquisado por meio do desenvolvimento de tecnologia de processamento natural da voz para avaliar linguagem oral e escrita, através de reconhecimento de fala e de análise sobre questões ligadas à estrutura ao discurso. Este tipo de tecnologia, complexa e cara, ainda está começando a ser desenvolvida e pesquisada.

2.2. Testes de proficiência oral e as novas tecnologias

As ferramentas necessárias para se construir um teste em língua estrangeira dependem do seu propósito e de questões práticas, tais como custo e disponibilidade de pessoas qualificadas e de tecnologia. Segundo Chapelle e Douglas (2006), o que temos hoje em dia, em termos de tecnologia desenvolvida para a área de avaliação são: sistemas modestos, planos de sistemas ideais e trabalhos em progresso.

Os autores citam questões básicas a serem pensadas em relação aos sistemas e arquitetura de uma ferramenta para o propósito da avaliação em língua estrangeira: processo de seleção da atividade (decisões sobre as tarefas, como: quantas?, quais?, por quê?), processo de apresentação (decisões sobre instruções e insumo), processo de resposta (como o candidato responderá o teste), processo de classificação do candidato (*scoring*).

No caso de um teste de proficiência oral, as questões tecnológicas trazem desafios ainda maiores, pela necessidade de desenvolvimento de programas de comunicação por voz, especialmente desenhados, com funções e recursos específicos, que possam ir ao encontro das demandas do teste.

Não há ainda notícia de teste de proficiência oral feito por meio de recursos de áudio e voz que comunicadores instantâneos (como Skype, MSN *Messenger*) oferecem pela Internet. Por enquanto, a tecnologia incorporada a este tipo de teste foi somente a do telefone, como é o caso da ACTFL-OPI, que pode ser feita com hora marcada.

Outro exemplo do uso desta tecnologia para fins de avaliação de proficiência oral automatizada é o PhonePass, citado por Chapelle e Douglas (2006), agora chamado SET-10, desenvolvido pela empresa Ordinate, Inc. (Ordinate Corporation, 2002b). Tal programa faz uso de uma nova tecnologia denominada NLP – em inglês, *natural language processing* – que estuda o processamento natural da fala. Os candidatos recebem as instruções impressas e ligam para uma central que os conecta ao computador no qual farão o teste para receber outras informações. Eles devem fornecer as respostas pelo telefone, e estas serão avaliadas pelo computador. O teste não tem o formato de entrevista, como na maioria dos testes de proficiência oral, mas contém uma variedade de tipos de tarefas, incluindo leitura de sentenças, repetição de sentenças, pronúncia de palavras disparadas por opostos semânticos e/ou por questões.

A avaliação é feita pelo próprio programa de computador segundo critérios que avaliam a acuidade na repetição de palavras, pronúncia, fluência de leitura e fluência de repetição. Esse programa é baseado em um sistema de reconhecimento de fala pelo computador que contém um algoritmo derivado de um grande corpus de língua falada de falantes nativos de diferentes regiões e dialetos sociais ingleses. O programa compara o desempenho do candidato com um molde que possui o vocabulário elicitado, características fonéticas das vogais e consoantes, assim como formulações e tempo de duração das sentenças.

Segundo seus criadores, *Ordinate Cooperation* (apud CHAPELLE; DOUGLAS, 2006), o dispositivo de avaliação eletrônica trabalha sem intervenção humana, somente sendo monitorado pelos administradores do programa e foi estudado extensivamente, tendo sido mostrada correlação de .94 com examinadores humanos para escores gerais e entre .79 e .89 para os escores dos subtestes.

Chapelle e Douglas (2006) citam ainda outras instituições que também estão desenvolvendo dispositivos para ajudar no trabalho dos examinadores, como por exemplo, a *Purdue University*, com o teste “Oral English Proficiency Test”. Eles estão trabalhando em um sistema de avaliação que automaticamente monitora o trabalho dos examinadores humanos e aponta itens problemáticos, mostrando marcas comparativas ou de referência (*benchmarks*). A intenção, segundo os pesquisadores, é manter o padrão estabelecido por aqueles que desenvolveram o teste nas avaliações dos examinadores.

De modo similar, no *Computerized Oral Proficiency Instrument* (COPI), desenvolvido pelo *Center of Applied Linguistics* em Hong Kong, a atribuição de nota é mediada por computador. Os examinadores escutam as respostas dos candidatos on-line, automaticamente tabulam os escores por meio de um algoritmo e armazenam as respostas dos candidatos e os escores.

Segundo os autores, o potencial deste tipo de dispositivo está ligado ao aumento de praticidade na aplicação dos testes. Entretanto, avisam que mais pesquisas são necessárias para se entender os usos e as limitações da tecnologia de processamento da fala natural (NLP), e que nosso desafio será compreender sob quais circunstâncias esses dispositivos são mais úteis do que os humanos para avaliar a produção linguística dos candidatos.

Dispositivos de comunicação síncrona, como o *MSN Messenger* e o *Skype*, tendem também a modificar as formas de avaliação, trazendo possibilidades de interação quase “presencial”, devido à utilização do vídeo e áudio em tempo real. Esse tipo de interação pode criar novos contextos para a avaliação de proficiência oral, que devem ser examinados, pois também estão suscetíveis a dificuldades e envolvem questões sobre construto e validade dos testes.

Chapelle e Douglas (2006) fazem uma previsão interessante e ousada para o futuro da avaliação em meio ao desenvolvimento tecnológico: a de que os computadores serão capazes de fazer uso do insumo falado de maneira adaptativa, emulando, assim, características de interação oral humana de maneiras ainda inimagináveis.

3. DISCUSSÃO DOS QUESTIONAMENTOS LEVANTADOS

Alderson (1990, apud CHAPELLE; DOUGLAS, 2006) delimitou algumas das mudanças que sugerem uma agenda inovadora a ser explorada pelo uso de computadores na área de testes de línguas: o computador consegue medir o tempo que o candidato leva para terminar uma tarefa, ou até mesmo para completar diferentes partes de uma tarefa, assim como consegue gravar informações sobre a rota seguida por ele através do teste. Além disso, o computador oferece diferentes maneiras de se apresentar informação, possibilitando diferentes tipos de insumo em formas variadas. Para o autor, computadores oferecem àqueles que desenvolvem os testes provisão de insumo rico multimodal na forma de vídeo, texto,

som e gráficos coloridos, aumentando potencialmente a autenticidade de ambos, o insumo e a resposta.

Baseando-nos, então, nos aspectos até aqui levantados, especialmente na análise de Chapelle e Douglas (2006) sobre as diferenças entre testes presenciais (papel e caneta ou entrevista) e testes por meio do computador (mediados por computador), e em nossa experiência enquanto participante e pesquisadora de um projeto de ensino e aprendizagem em meio virtual (“Teletandem Brasil: línguas estrangeiras para todos”), passamos agora a discutir as questões propostas no início deste artigo. Nossa intenção não é trazer respostas prontas que resolvam todos os problemas, mas sim levantar uma discussão acerca do assunto, baseada em nossas leituras e experiência com as novas tecnologias.

3.1. Que fatores devem ser levados em consideração ao desenvolver-se um teste de proficiência oral em contexto virtual, mediado por computador?

Retomamos aqui a primeira questão proposta e suas duas primeiras subperguntas: a) Que aspectos devem ser adequados para este meio? e b) Quais as possíveis vantagens e limitações de um teste de proficiência oral em tal contexto? Optamos por respondê-las conjuntamente, uma vez que os aspectos a serem adequados nos testes em contexto virtual implicam as vantagens e limitações trazidas pelos mesmos. Responderemos às perguntas baseando-nos nos seguintes fatores: circunstâncias econômicas, físicas e temporais, rubrica, fornecimento de insumo e autenticidade, e critérios de correção e construto.

3.1.1. Circunstâncias econômicas, físicas e temporais

Em relação aos aspectos econômicos, entendemos que a utilização de computadores para realização de testes de proficiência oral em língua estrangeira dispensa a utilização de papéis, fitas, rádios, tinta, gravadores. Mas, por outro lado, acrescenta possíveis gastos com cds, dvds e/ou bancos armazenadores de dados e ainda a construção e manutenção de páginas na Internet e programas de instruções.

Chapelle e Douglas (2006) citam como uma das limitações para realização de testes no computador a necessidade de se ter um bom equipamento e uma conexão (que deverá necessariamente ser rápida e eficiente). Devemos estar atentos ao fato de que os candidatos estão sujeitos a problemas tecnológicos com a máquina bem como a dificuldades com a conexão da Internet, que podem comprometer ou invalidar o desenvolvimento linguístico do candidato.

No caso de testes de proficiência oral, o uso de tecnologia de comunicação à distância traz vantagens quanto à conveniência de tempo e local, facilitando o acesso de pessoas em lugares distantes a qualquer hora, favorecendo, assim, os candidatos que se encontram longe do local onde o teste é aplicado. Entretanto, para que o teste se realize adequadamente, há a necessidade de bons equipamentos, adequados às exigências do teste.

Os equipamentos necessários apresentam limitações relacionadas a custos e manutenção, que podem encarecer o teste. A falta de padronização e necessidade de equipamento de configuração específica também pode causar limitações, tanto para aqueles que aplicam o teste quanto para os candidatos que o fazem. Entretanto, este tipo de

ferramenta tecnológica possibilita a gravação de testes e maior facilidade de armazenamento caso haja necessidade de revisão, correção posterior ao teste e/ou para treinamento de examinadores.

Não podemos nos esquecer também de fatores relacionados à segurança de um teste mediado por computador. O uso do computador gera flexibilidade e conveniência, mas ao mesmo tempo gera preocupações sobre a validade dos escores. Caso o teste seja feito pelo candidato em seu computador particular, ou qualquer outro que não esteja em um instituto de aplicação do teste, pode haver dificuldade em assegurar a real identidade do candidato, questão problemática especialmente em testes de grande repercussão. As pesquisas na área de tecnologia deverão cada vez mais enfatizar temas sobre a identificação do usuário.

Quanto aos examinadores, poderia haver mudanças na maneira como avaliam e na quantidade de examinadores presentes no teste. Ao invés de dois examinadores como comumente acontece nos testes presenciais, o candidato poderia ser avaliado por apenas um examinador, via recursos de áudio e vídeo de programas como mensageiros instantâneos, e o teste poderia ser gravado e (re)visto diversas vezes, pelo mesmo examinador ou por outro, em tempo real, antes de se atribuir a nota final ao candidato. O fato de o segundo examinador não estar presente no local do teste, mas assistindo-o via computador, também poderia ajudar a diminuir a pressão de ter duas pessoas dentro da sala avaliando o desempenho do candidato.

Deste modo, vê-se que ao se adaptar um teste de proficiência para o meio virtual surgem diversos questionamentos sobre como proceder, padronizar, agrupar e avaliar esses problemas tecnológicos com o equipamento e conexão de modo que nenhum candidato seja prejudicado ou beneficiado em função dessas questões.

Além dos gastos citados anteriormente com a manutenção de computadores, reforçamos a necessidade de um alto investimento inicial com a compra e instalação de equipamentos e de espaço e infraestrutura adequados para a instalação de computadores que possam ser usados pelos candidatos que não possuem o equipamento ou pelos examinadores. Tudo isso deverá ser totalmente planejado e bem estruturado, a fim de que diversas pessoas possam realizar o teste ao mesmo tempo e no mesmo ambiente, sem que prejudiquem involuntariamente o desempenho uns dos outros.

Sintetizamos no quadro a seguir as vantagens e limitações em relação às circunstâncias temporais, econômicas e físicas, discutidas até aqui, para instituições e candidatos, respectivamente.

	VANTAGENS	LIMITAÇÕES
Para instituições	<i>temporais</i>	
	maior acessibilidade para candidatos que se encontram longe dos locais de teste	necessidade de padronização de equipamentos
	<i>econômicas</i>	
	Possível diminuição de gastos com materiais impressos	necessidade de técnicos para instalação e manutenção de equipamentos
	<i>físicas</i>	
maior facilidade para armazenamento dos testes e revisão dos mesmos	necessidade de investimento em segurança (<i>high stakes tests</i> ²)	
Para candidatos	<i>temporais</i>	
	conveniência de tempo e local	conhecimento sobre uso do equipamento
	<i>econômicas</i>	
	gastos menores para locomoção	valor da inscrição do teste pode aumentar, devido a gastos com técnicos e manutenção de equipamentos
	<i>físicas</i>	
maior acessibilidade para candidatos que se encontram longe dos locais de teste	candidato necessitará ter bom equipamento e conexão de Internet (caso queira fazer o teste em casa)	

Quadro 1 – Vantagens e limitações das circunstâncias temporais, econômicas e físicas

Discutiremos na próxima subseção os fatores a serem levados em consideração ao desenvolver-se um teste de proficiência oral em contexto virtual em relação à rubrica do teste.

3.1.2. Rubrica – procedimentos e instruções para execução do teste

Quanto à rubrica do teste, as vantagens do uso do computador, segundo Chapelle e Douglas (2006), estão na possibilidade de dar ao candidato maior clareza nas instruções, inclusive com a opção de instruções em várias línguas, assim como maior oferecimento de ajuda durante o teste. Esta é uma vantagem em se tratando de uma nova tecnologia ainda não totalmente dominada por todos. Entretanto, esse tipo de ajuda pode retirar a uniformidade de execução do teste e dar mais chances àqueles que sabem utilizar melhor este tipo de tutoriais. Além disso, a criação de instruções e tutoriais para computadores demanda maior trabalho e conhecimento técnico por parte dos profissionais de diferentes áreas (como a ciência da computação e ensino de línguas, por exemplo).

A maior consistência para cronometrar o teste e suas partes também pode ser vista como vantagens do uso deste tipo de tecnologia, além da maior facilidade de provisão de materiais on-line ou por CD-ROM para preparação dos candidatos.

Considerando que, muitas vezes, o ambiente influencia a performance dos candidatos, aqueles que realizarem o teste em suas próprias residências possivelmente poderão se sentir mais à vontade e em uma situação mais confortável na hora da realização do exame do que aqueles que necessariamente precisam se deslocar até um local apropriado para fazê-

² *High stakes tests* são testes que provêm informação com base na qual decisões significativas serão tomadas sobre os candidatos, como, por exemplo, admissão em cursos ou em empregos (McNamara, 2000, p. 133).

lo. Isso denota uma despadronização quanto ao ambiente que pode produzir consequências diferentes (positivas ou negativas) no desempenho dos candidatos e, conseqüentemente, em sua nota final.

É importante reforçar que além do conhecimento dos tutoriais, a familiaridade do candidato com a máquina e seus programas e aplicativos pode interferir no desenvolvimento do teste e nos desempenhos, caso seja uma interação dinâmica e instantânea entre dois candidatos ou entre um candidato e um examinador. A possível falta de conhecimento sobre a utilização do equipamento pode comprometer o desempenho do candidato, assim como diferenças individuais podem alterar a avaliação do examinador.

Talvez fosse interessante, para uma maior viabilização na realização dos testes virtuais, a criação de um espaço virtual específico para este tipo de avaliação que se assemelhe aos aplicativos de comunicação instantânea como MSN *Messenger* e/ou *Skype*, com o diferencial de apresentar características específicas e funções apropriadas para este tipo de avaliação. Assim sendo, seria importante haver um primeiro contato dos candidatos com este tipo de aplicativo e seus recursos antes da realização do teste, pois o desconhecimento da ferramenta poderia prejudicar o desempenho, muitas vezes causando mais preocupações e ansiedade sobre o funcionamento da máquina do que com o teste em si.

Sintetizamos no quadro a seguir as vantagens e limitações em relação à rubrica dos testes em ambientes virtuais, discutidas até aqui, para instituições e candidatos, respectivamente.

	VANTAGENS	LIMITAÇÕES
Para instituições	<i>rubrica</i>	
	consistência e uniformidade na apresentação das instruções	diferentes níveis de familiaridade com a máquina e seus aplicativos podem retirar a uniformidade e alterar desempenhos
	possibilidade de se apresentar as instruções em várias línguas	
	maior consistência para cronometrar o teste como um todo e partes internas do teste	criação de instruções e tutoriais é mais trabalhosa do que em outros formatos de testes e envolve profissionais de outras áreas, como os da ciência da computação
facilidade de provisão de materiais para preparação e treino para o teste on-line ou por CD-ROM		
Para candidatos	<i>rubrica</i>	
	telas de ajuda – possibilidade de consulta de ajuda ou leitura de tutoriais durante a execução do teste	diferentes níveis de familiaridade com a máquina e seus aplicativos podem retirar a uniformidade e alterar desempenhos
Possibilidade de uso de materiais para preparação e treino para o teste on-line ou por CD-ROM		

Quadro 2 – Vantagens e limitações em relação à rubrica

Na próxima subseção discutiremos os fatores relacionados ao fornecimento de insumo e autenticidade nos testes de proficiência oral em ambiente digital.

3.1.3 Fornecimento de insumo e autenticidade

Para Chapelle e Douglas (2006), o computador possui ferramentas que podem gerar maior variedade de insumo em forma de informação visual e auditiva na composição das tarefas, contribuindo para a elaboração de um teste com melhor qualidade de diagramação e som e com custos menores, devido à possibilidade de uso de vídeo e gráficos e menor gasto de papel e tinta – embora haja outros tipos de gastos relacionados à manutenção de tais tecnologias, já citados anteriormente.

Acrescemos à qualidade e aos menores custos na composição das tarefas, o ambiente lúdico que determinadas ferramentas de mensagens síncronas instantâneas, como *MSN Messenger*, apresentam para conversações em áudio e vídeo e que ocasionam uma maior descontração para o aprendiz, podendo diminuir, conseqüentemente, a ansiedade e tensão próprias da situação de exame.

Tem-se ainda como vantagem, a possibilidade de uso de material autêntico, que pode gerar novos tipos de tarefas, como vídeos e figuras para começar uma discussão ou entrevista, advindos da própria Internet – embora deva haver cuidado quanto à fonte dos materiais e à possibilidade de rastreamento dos mesmos por sites de busca, pelos candidatos antes do teste.

Talvez alguns desafios estejam também relacionados a questões de autenticidade de um teste de proficiência oral via computador. Ao mesmo tempo em que a disponibilidade de maior variedade de insumo trazida por tais ferramentas pode contribuir para a geração de maior contextualização do material do teste, não podemos nos esquecer de pensar na verossimilhança com a realidade que o teste deve ter. A autenticidade, neste caso, pode ser vista sob duas perspectivas, a autenticidade interacional, que segundo Bachman (1991), é definida por meio da interação entre o conhecimento de língua do candidato e a tarefa comunicativa, e a autenticidade situacional, que segundo Chapelle e Douglas (2006), é definida pelas características do contexto, incluindo o cenário (local), os participantes, o conteúdo, o tom, o gênero. Deste modo, não basta que o material seja autêntico, ele deve ser também autenticado pelos tipos de tarefas propostos pelo teste e esses tipos de tarefas devem ser significativos para o candidato. No caso específico das interações virtuais síncronas, por meio de comunicadores instantâneos, acreditamos que a autenticidade situacional pode ser abalada caso o candidato não esteja familiarizado com o uso deste tipo de interação que simula virtualmente um contexto de comunicação face-a-face.

Há ainda restrições ligadas à disponibilidade da tecnologia necessária para o uso de alguns recursos, como os de transmissão de voz (VoIP), que estão sendo desenvolvidos e melhorados a cada dia, e também quanto ao uso que os candidatos normalmente fazem de tais tecnologias e sua familiaridade com tais recursos. Deste modo, a possível falta de familiaridade do candidato com os recursos gerados pelo computador e seus aplicativos e de verossimilhança com a realidade das tarefas propostas podem trazer conseqüências negativas advindas do efeito do método³ no desempenho do candidato e na interpretação dos resultados do teste.

³ Efeito do método (method effect) – segundo a teoria na área de avaliação em língua estrangeira, o método por meio do qual um teste é apresentado, as tarefas propostas, a maneira como os candidatos devem respondê-lo e os procedimentos usados para julgar o desempenho podem ter efeitos positivos e negativos sobre o desempenho dos candidatos (CHAPELLE; DOUGLAS, 2006, p.20).

Sintetizamos no quadro a seguir as vantagens e limitações em relação ao fornecimento de insumo e autenticidade dos testes em ambientes virtuais, discutidas até aqui, para instituições e candidatos, respectivamente.

	VANTAGENS	LIMITAÇÕES
Para instituições	<i>fornecimento de insumo e autenticidade</i>	
	maior possibilidade de variedade de insumo em forma de informação visual e auditiva na composição das tarefas	possíveis problemas em relação à falta de autenticidade, relacionados ao uso real que os candidatos normalmente fazem de tais tecnologias
Para candidatos	<i>fornecimento de insumo e autenticidade</i>	
	possível aumento de contextualização e autenticidade	possibilidade de consequências negativas advindas do efeito do método no desempenho do candidato e na interpretação dos resultados do teste

Quadro 3 – Vantagens e limitações ao fornecimento de insumo e autenticidade

Na próxima subseção abordaremos os fatores a serem levados em consideração em relação aos critérios de correção e construto de tais testes.

3.1.4 Critérios de correção e construto

De acordo com Chapelle e Douglas (2006), o uso de computadores para a área de avaliação pode trazer modificações em termos de critérios de correção e construto dos testes, ou seja, para a definição de quais aspectos do conhecimento ou habilidade do candidato estão sendo mensurados. Se por um lado, há a possibilidade de criação de novos tipos de tarefas, que poderão ser corrigidas por dispositivos eletrônicos, chamados *e-raters*, no caso específico dos testes de proficiência oral, a tecnologia necessária para tal façanha é mais complexa: a tecnologia de processamento natural da fala, que ainda é nova, cara e limitada.

Além disso, um teste de habilidade de produção oral ou escrita requer respostas linguisticamente construídas de possibilidades infinitas, e os dispositivos de correção automática ainda não estão bem desenvolvidos, podendo haver erros e comprometimento da validade do teste. Pode haver também um efeito retroativo negativo, fazendo com que os candidatos fiquem mais preocupados em aprender como o *e-rater* corrige o teste do que em estudar o conteúdo (CHAPELLE; DOUGLAS, 2006).

Por outro lado, a possível diminuição do número de entrevistadores ou mesmo a possibilidade de não estar na presença física deste, por meio da utilização de recursos de comunicação a distância mediados pelo computador, pode diminuir a carga de estresse para o candidato. Entretanto, há que se considerar os processos cognitivos envolvidos nas tarefas propostas e como a comunicação mediada por computador afeta o uso da linguagem, de modo que estes aspectos sejam contemplados no construto do teste. Segundo McNamara (2000), a diferença de formatos entre testes no papel e testes de desempenho, por exemplo, não é apenas uma coincidência, e sim reflete uma diferença implícita entre visões de linguagem e de uso da linguagem.

Deste modo, o desenvolvimento de testes de proficiência oral mediados por computador tende a trazer para a pesquisa em avaliação novos elementos para discussões acerca de definição de construto, critérios e procedimentos de avaliação, uma vez que tanto aqueles que desenvolvem tais testes quanto os implementadores (examinadores) deles devem compreender de que modo a dimensionalidade do conteúdo e as tarefas selecionadas refletem o modelo psicométrico ou sociolinguístico integrativo usado para determinar a relação entre o teste e os critérios de correção e procedimentos de avaliação. Ainda, reflexões sobre construto e critérios de correção poderão se respaldar na identificação de como os recursos tecnológicos afetam ou não o uso da linguagem pelo candidato e modificam a interpretabilidade dos escores.

Sintetizamos no quadro a seguir as vantagens e limitações discutidas nesta seção em relação aos critérios de correção e construto de um teste de proficiência oral em ambiente virtual para instituições e candidatos, respectivamente:

	VANTAGENS	LIMITAÇÕES
Para instituições	<i>critérios de correção e construto</i>	
	possibilidade de correção automática pelo computador, por meio do desenvolvimento de tecnologia de processamento natural da voz	tecnologia de processamento natural da fala é nova, cara e limitada
	possibilidade de geração de mais discussões e pesquisas acerca de definição de construto, critérios e procedimentos de avaliação ligados ao uso que o candidato faz da linguagem em tais ambientes	questões indefinidas sobre critérios de avaliação e correção podem gerar problemas em termos de construto e validade do teste
Para candidatos	<i>critérios de correção e construto</i>	
	possível diminuição do número de entrevistadores (ou a não presença física deste) pode diminuir a carga de estresse para o candidato	candidatos podem ficar mais preocupados em aprender como o <i>e-rater</i> corrige o teste do que em estudar o conteúdo

Quadro 4 – Vantagens e limitações dos critérios de correção e construto

Passaremos agora para a discussão da terceira subpergunta, que abrange o aspecto da validade em testes de proficiência oral mediados por computador.

3.2. O uso de computador neste caso pode ser inovador e ao mesmo tempo manter a validade que um teste de proficiência deve ter?

Para responder à terceira subpergunta, nos basearemos nas vantagens e limitações potenciais do uso do computador como tecnologia mediadora de um teste de proficiência oral, discutidos no item anterior, e nas idéias de Chapelle e Douglas (2006) com relação ao aspecto da validade em testes mediados por computador.

Segundo Chapelle e Douglas (2006), citando Wainer et al. (2000), a complexidade das questões relacionadas à validade em testes mediados por computador desafia abordagens simplistas ou prescritivas sobre validação:

Se o fato da apresentação dos itens em uma tela de computador, ao invés de em uma folha de papel, mudar os processos mentais exigidos para responder corretamente o item, a validade das inferências nestes escores pode ser mudada (WAINER et al, 2000 apud CHAPELLE; DOUGLAS, 2006, p.42).

Se por um lado, há necessidade dos pesquisadores entenderem melhor que aspectos da linguagem podem ser medidos por meio do computador, há, ao mesmo tempo, uma dificuldade, pois a própria tecnologia está mudando a maneira como a linguagem é usada. O uso de tecnologia para a avaliação, neste caso, pode gerar, então, maiores discussões e pesquisas acerca de aspectos do uso da linguagem no meio virtual.

A elaboração de tarefas para testes de proficiência oral mediados por computador deve ser feita levando-se em conta os processos mentais do candidato para a utilização dos recursos de comunicação à distância exigidos para este tipo de teste. Estudos sobre validade de testes em meio virtual poderiam, assim, enfocar o aspecto da validade por este viés.

Ainda, ao que parece, a questão da validade se centra também de maneira um tanto intensa no campo do uso correto da tecnologia e do impacto desse uso sobre as inferências feitas pelo examinador. A apreensão do candidato em relação ao computador pode afetar seu desempenho linguístico, a falta de familiaridade com as ferramentas pode trazer resultados para sua autoconfiança ou para o tempo que ele leva para completar o teste, afetando, também, seu desempenho. Há preocupação por parte daqueles que desenvolvem assim como dos que aplicam testes sobre uma possível mudança no desempenho do candidato por causa do novo formato de teste (mediado por computador).

Acreditamos também que questões de ansiedade originárias da falta de familiaridade com o(s) aplicativo(s) que permitem a interação entre candidato-candidato e/ou candidato-examinador, gravação, envio e armazenamento de dados e a possível apreensão causada por um primeiro contato com o parceiro de exame ou examinador de maneira virtual podem prejudicar o desempenho do candidato.

Essas inovações tecnológicas também são suscetíveis a falhas, que juntamente com problemas de conexão e identificação, podem contribuir para aumentar a ansiedade e tensão do candidato, que possivelmente será refletida em seu desempenho e, conseqüentemente, em sua nota final.

Entendemos, deste modo, que os testes de proficiência oral por meio da utilização de recursos tecnológicos estão sujeitos a falhas, à imprevisibilidade da máquina e da conexão, assim como em outros tipos de testes utilizando os mesmos recursos, que necessariamente devem ser estudados e levados em consideração no momento de elaboração das tarefas e do exame em si, para que não comprometam o desempenho dos examinandos e a validade e confiabilidade dos resultados.

4. CONSIDERAÇÕES FINAIS

Concordando com Chapelle e Douglas (2006), as mudanças nesse momento aparentam ser mais evolucionárias do que revolucionárias. O que parece, na verdade, é que as ameaças e preocupações que permeiam a pesquisa e uso da TLMC, principalmente para testes de

proficiência oral, são as mesmas existentes em outros formatos de testes, como os de papel e caneta, por exemplo.

Ao buscarmos um denominador comum dentre as limitações trazidas a respeito dos fatores a serem levados em consideração para a formulação de testes de proficiência oral em ambiente virtual (circunstâncias econômicas, físicas e temporais, rubrica, fornecimento de insumo, autenticidade, critérios de correção, construto e validade), percebe-se que elas parecem girar em torno de questionamentos sobre se e como os elementos tecnológicos podem afetar o uso da linguagem e o desempenho do candidato no teste.

Acreditamos que pesquisas enfocando as circunstâncias físicas e temporais relacionadas a tais testes poderiam estudar de que modo a presença física se difere da presença virtual (ou seja, mediada por computador) nos testes de proficiência oral e se e como o desempenho do candidato é afetado por tais elementos que causam a distinção entre elas. Pesquisas sobre a rubrica de testes de proficiência oral mediados por computador poderiam investigar se e de que modo os tutoriais on-line e a familiaridade com os recursos agregam ou retiram uniformidade na execução dos testes, ou, ainda, se como o desconhecimento do uso das ferramentas e aplicativos, preocupações e ansiedade sobre o funcionamento da máquina podem alterar o desempenho dos candidatos.

As investigações sobre o fornecimento de insumo e autenticidade poderiam avaliar de que modo a verossimilhança com a realidade na situação de comunicação mediada por computador (ou falta dela) e as tarefas propostas podem trazer consequências na interpretação dos resultados do teste. E, por fim, pesquisas sobre critérios de correção e construto poderiam enfocar de que modo o uso da linguagem oral difere na comunicação mediada por computador e como isso pode se refletir na construção dos testes de proficiência oral em ambiente virtual.

Mudanças revolucionárias para esta área deverão, assim, ser motivadas e apoiadas pelo entrecruzamento do avanço das tecnologias e de nosso entendimento sobre as questões levantadas neste trabalho.

REFERÊNCIAS BIBLIOGRÁFICAS

- ALDERSON, J.C.; CLAPHAM, C.; WALL, D. (1998). *Exámenes de idiomas. Elaboración y evaluación*. Trad. Neus Figueras. Edición española. Madrid: Cambridge University Press.
- BACHMAN, L.F. (1988). Problems in examining the validity of the ACTFL oral proficiency interview. *Studies in Second Language Acquisition*, n.10, p. 149–64.
- BACHMAN, L.F. (1991). What does language testing have to offer? *TESOL Quarterly*, v. 25, n.4, p. 671-704.
- BROWN, A. (2003). Interviewer variation and the co-construction of speaking proficiency. *Language Testing*, v.20, n.1, p.1–25.
- BRUNNER, J.J. (2004). Educação no encontro com as novas tecnologias. In: TEDESCO, J.C. (Org.). *Educação e novas tecnologias*. Trad. Claudia Berliner, Silvana Cobucci Leite. São Paulo: Cortez; Buenos Aires: Instituto Internacional de Planeamiento de la Educación; Brasília: UNESCO, p. 17-75.

- CHAPELLE, C.A.; DOUGLAS, D. (2006). *Assessing language through computer technology*. Cambridge: CUP.
- CONSOLO, D.A. (2004). A construção de um instrumento de avaliação da proficiência oral do professor de língua estrangeira. *Trabalhos em Linguística Aplicada*, Campinas, v.43, n.2, p. 265-286, jul./dez..
- FULCHER, G. (2003). The history of testing second language speaking. In: Fulcher, G. *Testing second language speaking*. Harlow: Pearson-Longman, p.1-17.
- MARTINS, T.H.B. (2005). *Subsídios para a elaboração de um exame de proficiência para professores de inglês*. Dissertação de Mestrado em Linguística Aplicada. Instituto de Estudos da Linguagem, Unicamp. Campinas.
- MCNAMARA, T. (2000). *Language Testing*. Oxford, OUP.
- WARSCHAUER, M.; HEALEY, D. (1998). Computers and language learning: an overview. *Language Teaching*, 31, p.57-71.
- WARSCHAUER, M. (2000). The changing global economy and the future of English teaching. *TESOL Quarterly*, v.34, n.3, p.511-535.
- WARSCHAUER, M. (2004). Technological change and the future of CALL. In: FOTOS, S.; BROWN, C. (Eds.). *New perspectives on CALL for Second and Foreign Language Classrooms*. Mahwah, NJ: Lawrence Erlbaum Associates, p. 15-25.

Recebido: 25/07/2010

Aceito: 26/11/2010

