

KÁTIA DA SILVA ARAÚJO

A PERSPECTIVA DO EXAMINANDO SOBRE A
AUTENTICIDADE DE AVALIAÇÕES EM LEITURA EM
LÍNGUA ESTRANGEIRA

Dissertação apresentada ao Curso de
Linguística Aplicada do Instituto de Estudos
da Linguagem da Universidade Estadual de
Campinas como requisito parcial para
obtenção do título de Mestre em Linguística
Aplicada na Área de Ensino-Aprendizagem de
2ª Língua e Língua Estrangeira.

Orientadora: Profa. Dra. Matilde Virgínia
Ricardi Scaramucci

Universidade Estadual de Campinas
Campinas, SP, Brasil.
2007

Ficha catalográfica elaborada pela Biblioteca do IEL - Unicamp

Ar15p	<p>Araújo, Kátia da Silva. A perspectiva do examinador sobre a autenticidade de avaliação em leitura em língua estrangeira / Kátia da Silva Araújo. -- Campinas, SP : [s.n.], 2007.</p> <p>Orientador : Matilde Virgínia Ricardi Scaramucci. Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Estudos da Linguagem.</p> <p>1. Avaliação. 2. Leitura. 3. Língua estrangeira. I. Scaramucci, Matilde Virgínia Ricardi. II. Universidade Estadual de Campinas. Instituto de Estudos da Linguagem. III. Título.</p>
-------	---

Título em inglês: Test takers' perspective of authenticity in foreign language reading assessment.

Palavras-chaves em inglês (Keywords): Assessment; Reading; Foreign language.

Área de concentração: Língua estrangeira.

Titulação: Mestre em Lingüística Aplicada.

Banca examinadora: Profa. Dra. Matilde Virgínia Ricardi Scaramucci (orientadora), Prof. Dr. Paulo Sampaio Xavier de Oliveira e Prof. Dr. Hécio de Pádua Lanzoni.

Data da defesa: 30/01/2007.

Programa de Pós-Graduação: Programa de Pós-Graduação em Lingüística Aplicada.

AGRADECIMENTOS

Aos pesquisadores que colaboraram com a pesquisa, sem os quais este trabalho não seria possível.

BANCA EXAMINADORA

A professora Matilde Scaramucci, pela dedicação, competência e respeito com que me recebeu.

Aos professores Denise Bertoli Braga, Joanne Marie McCaffrey Bastardo Netto, Paulo Oliveira e Hécio de Pádua Lanzoni, pelas sugestões, críticas e orientações que ajudaram a corrigir este trabalho.

A CAPES e à Fapesp, pelo apoio financeiro recebido.

As minhas amigas Davi, Sábina, Lítiane, Ricardo, Lívia, Tânia, Renata, Nárciso, Ana Paula, Brézilda, Emília e Walter, por me incentivarem e distraírem os momentos de estresse, e também por serem capazes de fazer a distância.

Prof. Dra. Matilde Virgínia Ricardi Scaramucci

As meus amigos Davi, Sábina, Lítiane, Ricardo, Lívia, Tânia, Renata, Nárciso, Ana Paula, Brézilda, Emília e Walter, por me incentivarem e distraírem os momentos de estresse, e também por serem capazes de fazer a distância.

Prof. Dr. Paulo Sampaio Xavier de Oliveira

Prof. Dr. Hécio de Pádua Lanzoni

200731805

AGRADECIMENTOS

Aos examinandos que colaboraram com a pesquisa, sem os quais esse trabalho não seria possível.

À professora Matilde Scaramucci, pela dedicação, competência e respeito com que me orientou.

Aos professores Denise Bértoli Braga, Joanne Marie McCaffrey Busnardo Neto, Paulo Oliveira e Hécio de Pádua Lanzoni, pelas sugestões, críticas e orientações que ajudaram a compor este trabalho.

À CAPES e à Fapesp, pelo apoio financeiro recebido.

Às minhas companheiras de casa Juliana, Mariza, Eliza, Cláudia e Cris, pelos estresses e alegrias compartilhados.

Aos meus amigos Davis, Rubiana, Liliane, Ricardo, Lívia, Thaís, Renata, Nazaré, Ana Paula, Fernanda, Estefânia e Walter, por me incentivarem e distraírem na medida certa, e algumas vezes pela proximidade apesar da distância.

À Viviane e ao Ronaldo, pelo carinho que só os irmãos podem dar.

E aos meus pais, Eulália e Joaquim, por tudo que me ensinaram – coisas que eu nunca encontraria nos livros e nem poderia ter aprendido melhor com mais ninguém.

“Se estamos aqui reunidos estou contente. Penso com alegria que tudo quanto escrevi e vivi serviu para nos aproximar. É o primeiro dever do humanista e a fundamental tarefa da inteligência assegurar o conhecimento e o entendimento entre os homens. Bem vale haver lutado e cantado, bem vale haver vivido se o amor me acompanha.”

De Para nacer he nacido, Pablo Neruda.

RESUMO

A “autenticidade” de uma avaliação de língua estrangeira (LE) geralmente é uma característica tida como positiva e desejável, pois refletiria princípios mais condizentes com o ideal de avaliação de desempenho comunicativo. A “autenticidade” se caracterizaria pelo uso de insumo “real” (não produzidos especialmente para a avaliação), e tarefas integradoras, que se aproximem dos usos cotidianos da LE. A perspectiva dos examinandos revela-se importante nesse contexto, uma vez que suas concepções acerca do que são, por exemplo, leitura e avaliação, orientarão suas percepções sobre um determinado exame, o modo como o realizam e seu conseqüente desempenho. Busca-se aqui compreender melhor qual é a percepção do examinando sobre a “autenticidade” do teste – quais elementos ele identifica como “autênticos” em três tipos de avaliação de leitura, qual a importância atribuída a essa autenticidade e se suas concepções são consoantes com princípios mais (ou menos) integradores. Para isso, metodologia qualitativa é empregada, com estudos de caso que permitem refletir sobre a natureza dessas percepções. Os resultados refletem diferenças nas percepções sobre autenticidade entre examinandos com diferentes níveis de proficiência, assim como na importância atribuída à autenticidade e à integração de habilidades na avaliação.

Palavras-chave: Avaliação, Leitura, Língua estrangeira.

ABSTRACT

The “authenticity” of a foreign language (FL) assessment generally is taken as a positive characteristic, since it would reflect principles more suitable to the idea of communicative performance assessment. “Authenticity” could be described as the use of authentic input (not produced only for assessment purposes) and integrated tasks, which would be similar to the common FL use. The perspectives of examinees are important in this context, since their conceptions about reading and assessment, for instance, will lead their perceptions about an exam, the manner they do it and their consequent performance. It is aimed to comprehend examinees’ perceptions of authenticity of tests – which elements they identify as authentic in three kinds of reading assessment, the importance given to authenticity, and if their conceptions are consonant with more(or less) integrated principles. To achieve that, qualitative methodology is used, with case studies that permit to reflect about the essence of these perceptions. The results demonstrate differences in the perceptions of authenticity among examinees with different levels of proficiency, as well as in the importance given to authenticity and to the integration of skills in assessment.

Key words: Assessment, Reading, Foreign Language.

SUMÁRIO

CAPÍTULO I : Introdução	
1.1.Introdução	15
1.2.Justificativa	17
1.3.Problema e questão de pesquisa.....	20
1.4.Metodologia	21
1.4.1. Instrumento de coleta de dados: entrevistas semi-estruturadas	24
1.4.2. Materiais de coleta de dados: avaliações com diferentes níveis de integração de habilidades	25
1.4.3. Sujeitos	26
1.5. Estrutura da dissertação	28
 CAPÍTULO II : Fundamentação teórica.....	 31
2.1. Integração e isolamento de habilidades	33
2.1.1. Surgimento e evolução do conceito de integração.....	33
2.1.2. A atual discussão acerca da integração e isolamento de habilidades	40
2.1.3. Integração de habilidades nos exames de proficiência	43
2.2. Autenticidade na avaliação	49
2.2.1. A autenticidade e os “testes diretos”	50
2.2.2. A relatividade da autenticidade	53
2.3. A perspectiva do examinando.....	56
2.3.1. A validade de face	58
2.3.2. Diferentes contextos da autenticidade na avaliação	61
 CAPÍTULO III: Análise e discussão dos dados	 67
3.1. Estudo piloto	67
3.2. Coleta de dados.....	76
3.2.1. Caracterização das tarefas.....	70
3.2.1.1. Tarefa 1: Múltipla escolha.....	78

3.2.1.2. Tarefa 2: Resumo.....	80
3.2.1.3. Tarefa 3: E-mail.....	82
3.2.2. Procedimentos de coleta de dados	84
3.2.3. Perfis dos sujeitos	87
3.3. Análise dos dados	91
3.3.1. Conhecimento de vocabulário e de gramática	92
3.3.2. Integração de habilidades na avaliação.....	93
3.3.3. Temas dos textos usados nas tarefas.....	96
3.3.4. Tarefa 1	99
3.3.5. Tarefa 2	105
3.3.6. Tarefa 3	113
CAPÍTULO IV: Conclusões.....	125
REFERÊNCIAS BIBLIOGRÁFICAS	135
APÊNDICES	139
APÊNDICE I: Roteiro de aplicação da entrevista	139
APÊNDICE II: Teste Cloze	140
APÊNDICE III: Materiais de coleta de dados	142
APÊNDICE III.I : Tarefa 1.....	142
APÊNDICE III.II : Tarefa 2	144
APÊNDICE III.III : Tarefa 3	146

CAPÍTULO I

A TRAJETÓRIA DA PESQUISA

1.1. Introdução

A expressão “uso real de língua” é constantemente utilizada nos estudos sobre ensino/aprendizagem de Língua Estrangeira (doravante LE) e pelas instituições que realizam esse ensino quando se referindo a uma visão mais voltada para uma vertente comunicativa desse processo. Na avaliação de línguas, a integração de macro-habilidades (leitura avaliada junto com escrita, por exemplo) é um dos aspectos relacionados à idéia de “uso real de língua” ou “autenticidade” pois, quando pensamos em uso de língua, uma das primeiras idéias que surgem é a da simultaneidade do ler, escrever, falar e compreender a fala como reflexo desse uso. Dessa forma, se o objetivo de uma avaliação de língua é tentar ser uma amostra representativa do uso da linguagem em situações reais – o que nos remete ao conceito de validade da avaliação –, o mais coerente seria formular itens que reflitam essa concepção de linguagem, que integrem o uso de macro-habilidades (leitura, escrita, oralidade e compreensão auditiva) na tentativa de aproximar o instrumento avaliativo do tipo de desempenho que se espera que o examinando¹ apresente.

Operacionalizar essas idéias em forma de exames, no entanto, revela-se mais complexo e dependente de outros fatores, como a relatividade do conceito de “autenticidade”. Por ser um conceito relativo a quem são os participantes envolvidos, que usos realizam da LE e com que propósitos o fazem, às experiências de aprendizagem e avaliação que trazem, entre outros fatores, não podemos tomar um teste como “autêntico por natureza” sem analisar a quem é dirigido, quais são os objetivos avaliativos, o contexto de aplicação etc. Da mesma forma, não se pode dizer que a integração de macro-habilidades, que muitas vezes é tomada como um fator que tornaria avaliações de LE mais “autênticas”, o seja de fato, independente do contexto em que se insere. Este estudo busca investigar alguns aspectos presentes nessa relatividade, refletindo sobre como ela se

¹ Usaremos neste trabalho o termo *examinando* como equivalente ao termo *test taker*, largamente empregado nos estudos reportados em inglês sobre avaliação de língua, referindo-se ao sujeito que é submetido a uma avaliação de línguas.

configura nas percepções de diferente examinandos frente a exames com diferentes níveis de integração.

Basicamente, três tópicos estão envolvidos no problema de pesquisa no presente estudo: *i)* o conceito de integração de itens avaliativos, *ii)* autenticidade na avaliação e *iii)* perspectiva dos examinandos. A integração de habilidades será tomada segundo o viés da autenticidade na avaliação de LE, dada a natureza da discussão que envolve essa integração na avaliação de línguas, que versa sobre visões de linguagem e o modo como elas se refletem nos instrumentos avaliativos. Daí a necessidade de se investigar a perspectiva dos examinandos – já que a autenticidade se constitui como característica relativa, sujeita a flutuações de grau e natureza, não absoluta – e essas percepções serem determinantes na construção de instrumentos avaliativos válidos do ponto de vista da autenticidade.

As hipóteses iniciais desse problema de pesquisa versam sobre a relatividade da autenticidade na avaliação. Partindo do princípio que não existe um conceito absoluto de autenticidade, busca-se investigar se a visão de autenticidade trazida pelos sujeitos, quando/se identificada em algum dos itens, constituiria traço significativo para atribuição de validade de face desse item. A validade atribuída pelo leigo é o que na literatura da área de avaliação de línguas é denominada *validade de face* – um tipo de validade “aparente”, que diz se o item *parece* ser válido, ou seja, se parece ser um instrumento adequado ao que propõe avaliar, baseado apenas em seu aspecto mais superficial, sem a análise técnica empreendida dentro dos rigores necessários para a validação feita por um especialista da área. Este conceito será discutido de forma mais detalhada no Capítulo II.

Por ser o tipo de validade atribuída pelo leigo, a validade de face muitas vezes é tratada como uma característica “menor” de um teste. No entanto, sua essência diz respeito a aspectos importantes num exame, como a aceitação que ele tem socialmente pode-se dizer que as percepções que um examinando tem sobre a validade de um exame influenciam o crédito que ele confere a ele. Embora possa parecer uma simples questão de “relações públicas”, ao ampliar um pouco mais a discussão, poderíamos dizer que a validade de face influi no modo como examinandos acreditam que devam proceder durante o teste – que conhecimentos e capacidades mobilizar, que estratégias ativar, que tipo de preparação empreender etc.

Assim, o que orienta o trabalho aqui proposto é uma indagação sobre como se configura o conceito de autenticidade na avaliação para aprendizes de inglês-LE e sobre a influência dessa idéia de autenticidade na constituição da validade de face dessas avaliações.

Para responder a esses questionamentos, metodologia de caráter qualitativo foi empregada no trabalho. O instrumento de coleta de dados usado foi a entrevista semi-estruturada e os materiais utilizados foram três tipos diferentes de avaliações de leitura em LE. A escolha desses instrumentos e materiais são justificadas e descritas em *1.4 Metodologia* e em *3.2 Coleta de dados*, onde relato o processo de seleção e construção dos materiais, os princípios que orientaram o estudo piloto e como ele contribuiu para esse refinamento metodológico. No mesmo sentido, o perfil requerido dos sujeitos de pesquisa também sofreu ajustes depois da realização do estudo piloto e, para a coleta efetiva de dados, foram escolhidos sujeitos com perfis que representassem diferentes necessidades e vivências com relação à leitura em LE e a avaliações de LE. Tomando a autenticidade das avaliações como relativa, dependente do tipo de contexto e situação de uso de língua a que estamos nos referindo para definir o que é ou não autêntico, natural foi buscar contemplar diferentes tipos de perfis de sujeitos, que poderiam expressar diferentes percepções sobre a autenticidade.

1.2. Justificativa

A idéia de considerar o *saber usar* a língua como construto passível de avaliação, em detrimento do *saber sobre a língua* é uma das grandes mudanças que a área de avaliação de línguas viu gradativamente desenvolver-se desde os primeiros empreendimentos teóricos e práticos nesse campo até as pesquisas mais recentes.

Essas mudanças influenciam não apenas contextos diretamente ligados a algum tipo de avaliação de línguas, como uma sala de aula, mas têm um alcance mais amplo. Em primeiro lugar, elas são parte de um complexo muito mais vasto: essas “evoluções” de conceitos e práticas não são solitárias, mas imbricadas e acompanhadas, em diferentes ritmos, por mudanças em toda a área de ensino e aprendizagem de línguas, de visões sobre linguagem e educação e das efetivas práticas relativas a esses campos. Pensar sobre

avaliação de línguas, portanto, significa refletir de modo mais apurado sobre um dos componentes de um amplo enredamento social e educacional. Paralelamente, outro fator que corrobora o alcance amplo do estudo da avaliação de línguas é o papel que ela desempenha na sociedade num sentido mais “externo” ao da Lingüística Aplicada, no de avaliações de línguas como instrumento utilizado em diversos segmentos sociais. A pesquisa aqui realizada visa a fornecer maior entendimento desse objeto para que ganhos sejam trazidos para o modo como é usado, em todas suas etapas — elaboração, aplicação, correção. Esse uso mais competente tem implicações que abrangem não apenas o âmbito da sala de aula de língua, mas também outras esferas que necessitam da aferição de conhecimento, capacidades ou habilidades em uma LE. Em muitas situações, as inferências feitas sobre as capacidades lingüísticas dos examinandos são usadas na tomada de decisões que envolvem esses indivíduos ou instituições às quais estão ligados. Um indivíduo pode ser considerado inapto para uma posição ou cargo que esteja pleiteando, ou um currículo escolar ser modificado baseado nos resultados de exames de larga escala (como o ENEM [*Exame Nacional do Ensino Médio*], SAEB [*Sistema Nacional de Avaliação da Educação Básica*]² etc.). Dessa forma, quanto mais claros forem os princípios, propósitos e métodos sobre os quais se fundamentam as avaliações de línguas, melhores e mais conscientes poderão ser as inferências sobre a competência lingüística dos examinandos, assim como as decisões baseadas nelas (cf. Bachman & Palmer 1996, p. 95). A investigação sobre a avaliação de compreensão em leitura em LE, objeto de estudo aqui, foi tomada como tópico de pesquisa a partir dessas reflexões, e busca desenvolver-se segundo esses princípios.

Para citar um exemplo que por si só já é abrangente, temos os certificados de proficiência em LE, exigidos por agências de fomento à pesquisa para concessão de bolsas de estudo no exterior, por universidades que aceitam alunos estrangeiros, para o deferimento de visto de trabalho para estrangeiros, entre outros usos. Além desses, deve-se atentar para o âmbito educacional mais estrito — o das *avaliações de rendimento*, não vinculado a avaliações externas, como certificados de proficiência ou exames vestibulares, mas que envolve mais especificamente a prática de sala de aula de Língua Estrangeira. A

² Esses são alguns exemplos de exames oficiais do governo brasileiro, de cujos resultados dependem, a curto, médio e/ou longo prazo, os indivíduos que se submetem a eles e a própria Educação do país como instituição social.

sala de aula é o espaço clássico de ocorrência da avaliação e, mesmo assim (ou, talvez, justamente por isso), é nesse ambiente em que se pode encontrar o maior número de testes³ mal elaborados, falta de critérios claros de correção, julgamentos de valor equivocados devido ao desconhecimento sobre o uso apropriado do instrumento avaliativo. Muitas vezes, a visão de avaliação como instrumento punitivo ou de “prova de conhecimento” é sobreposta ao caráter de instrumento que busca aferir a capacidade de uso da língua em diferentes contextos.

Apenas o uso da expressão “uso da língua em diferentes contextos” já denota uma certa concepção de língua e avaliação – a de língua vista como um construto vivo, passível de ser “colocado em prática”, utilizado com uma certa finalidade, determinada pelo universo de produção e recepção. A avaliação de língua, segundo essa concepção, poderia ser definida como um instrumento que busca fornecer um quadro sobre a capacidade do examinando em usar a língua em diferentes situações. Essa definição se opõe àquela calcada numa visão mais estruturalista de língua, em que ela é considerada mais como um objeto de “descrição” do que de “uso”. A avaliação inspirada por esse tipo de concepção não serviria ao propósito de funcionar com amostragem da habilidade comunicativa de um sujeito numa língua, mas apenas do quanto ele sabe *sobre* a língua.

É a partir dessa reflexão que este trabalho tem seu ponto de partida. A questão da *autenticidade*, do quanto os testes de língua refletem o uso que é feito dela na vida real, é um dos cerne da discussão sobre a utilização de itens avaliativos *integradores*, que integram o uso de diferentes habilidades num mesmo item, opostos aos itens mais *isolados* ou *independentes*, que buscam isolar um determinado elemento, ou habilidade, como objeto de avaliação. Na avaliação de leitura em LE, notamos uma grande necessidade de reflexão sobre esse tópico – como afirma Wielewicki (1997), poucas são as pesquisas sobre a avaliação da leitura em contexto brasileiro, e muitos são os fatores que contribuem para confirmar a importância da pesquisa sobre a avaliação desse construto.

³ O termo *teste* é usado neste trabalho como sinônimo de *avaliação* e *exame*.

1.3. Problema e questão de pesquisa

Como já expus anteriormente, os três tópicos envolvidos no problema de pesquisa no presente estudo são *i)* o conceito de integração de habilidades na avaliação, *ii)* autenticidade na avaliação e *iii)* perspectiva dos examinandos. A integração de habilidades está relacionada à autenticidade na medida em que o princípio básico da integração é buscar semelhança com a língua que é usada na vida real, *autenticamente* – a língua não é usada na vida real em unidades independentes de um todo, ou mesmo macro-habilidades isoladas de contexto. A noção de diferentes perspectivas vem somar-se a esse arcabouço devido à natureza da idéia de “autenticidade”. Diferentes pessoas, com diferentes vidas e diferentes “usos reais de língua”, demonstrarão diferentes percepções sobre como esses usos são representados nas avaliações. Ou seja, a autenticidade é relativa a essas percepções e, por isso, não se pode falar dela como uma característica absoluta, presente ou não nos testes. O mais adequado seria falar se num determinado contexto, para um determinado grupo de pessoas, aquele teste representa uso real de língua e se elas o percebem como “autêntico”.

Assim, partindo do princípio que não existe um conceito absoluto de autenticidade, busca-se investigar se a visão de autenticidade trazida pelos sujeitos, quando/se identificada em algum dos itens, constituiria traço significativo para atribuição de validade de face desse item.

Isso posto, pode-se dizer que a pergunta de pesquisa é: **como se configura o conceito de autenticidade na avaliação de leitura em LE para os examinandos?** Busca-se tanto investigar como eles percebem a validade de face dos instrumentos – isto é, qual é a percepção sobre a aparente validade do item, se para eles (examinandos) o item parece avaliar o construto a que se propõe – quanto o que “autenticidade” da avaliação significa e como/se ela influencia o processo de realização dos testes.

Dadas essas diretrizes de trabalho, foi empregada aqui metodologia de caráter qualitativo. Uma descrição mais detalhada acerca do método de coleta de dados é feita na seção 1.4. *Metodologia*, onde é discutido o desenho metodológico da pesquisa, os materiais e instrumentos utilizados, os perfis dos sujeitos e as justificativas dessas escolhas. Esses instrumentos, materiais e sujeitos serão caracterizados mais detalhadamente na descrição da coleta de dados, em 3.2. *Coleta de dados*.

1.4. Metodologia

O uso de metodologia qualitativa na pesquisa em avaliação em LE vem ganhando mais espaço na área recentemente (ver Banerjee & Luoma, 1997), embora nos países de mais tradição em pesquisas nesse campo, majoritariamente falantes de inglês, os métodos quantitativos ainda sejam maioria – basta uma rápida visita aos artigos publicados em *Language Testing* e *Language Testing Update*, periódicos mais importantes da área de avaliação de línguas, para averiguar esse fato. Os métodos qualitativos, entretanto, vêm ganhando adeptos que buscam conferir natureza diferente aos seus estudos, aprofundando suas análises e investigando nuances dos dados que métodos quantitativos dificilmente alcançariam.

O propósito almejado com a escolha de metodologia qualitativa neste trabalho foi justamente este: buscar ir mais a fundo nessas investigações, embora com um número mais reduzido de sujeitos. Essa escolha não é aleatória, mas orientada pelo propósito de pesquisa. Como salientam Larsen-Freeman & Long (1991, p. 14) é a questão de pesquisa que orienta a escolha do desenho metodológico. A investigação de percepções através de metodologia quantitativa é uma escolha viável (v. Scaramucci 2004c, Bartholomeu 2002), mas, no presente estudo, a abordagem qualitativa se fez mais adequada.

O uso dessa metodologia veio responder à necessidade aqui presente de um tipo de análise mais afim à natureza dos dados que se esperava recolher. O paradigma interpretativista de pesquisa visa primeiramente à leitura da realidade através de instrumentos que levem em consideração a natureza desses objetos, cujos significados não se encontram prontos como fenômenos físicos exatos, mas são atribuídos pelas pessoas. O lidar com os dados é feito de maneira indireta, através da interpretação desses vários significados (cf. Moita Lopes, 1994). Além disso, como apontam Lüdke e André (1986, p. 13), citando Bogdan & Biklen (1982), *a pesquisa qualitativa (...) envolve a obtenção de dados descritivos, obtidos no contato direto do pesquisador com a situação estudada, enfatiza mais o processo do que o produto e se preocupa em retratar a perspectiva dos participantes*. Tudo isso se conforma ao problema de pesquisa aqui estabelecido, à natureza das indagações feitas e ao modo como se mostra mais adequado dedicar-se a esses dados e análises. Dessa forma, como já disse anteriormente, embora o uso de metodologia

quantitativa seja possível (e efetivamente empregado em estudos anteriores na investigação de percepções), o cunho qualitativo mostrou-se mais atraente devido à natureza da pergunta de pesquisa, que busca refletir sobre *como se configura* um conceito segundo as perspectivas de um determinado grupo.

Um outro ponto que orienta o uso de metodologia qualitativa no trabalho é justamente o mencionado na citação acima de Bogdan & Biklen (*apud* Lüdke e André 1986, p. 13) sobre as noções de foco no produto ou no processo. Isso também é ressaltado por Banerjee & Luoma (*op.cit.*, p. 276), ao afirmarem que métodos qualitativos de investigação trazem mais esclarecimentos sobre a natureza dos conteúdos avaliados e dos desempenhos sobre os quais as notas são baseadas, assim como dos processos envolvidos no realizar do teste. E investigar as percepções que examinandos têm sobre avaliações de língua é um propósito de pesquisa mais voltado para o *processo* de realização dos testes. Uso o termo “mais voltado” porque não considero prudente dizer que é “totalmente” voltado para o processo. Uma dicotomia entre processo e produto geralmente é estabelecida nas pesquisas, mas assim como para Kaplan (1988) e para Scaramucci (1995, p. 99-106), acredito que não se trata de termos dicotômicos, mas que se completam: todo processo resultará em produto, e todo produto é resultado de algum processo. Aplicando essa fórmula para a área da avaliação de línguas, pode-se dizer que o processo de realizar um teste não pode ser desvinculado do desempenho efetivamente alcançado, e esse desempenho é resultado da atuação de vários componentes durante o processo de realização do teste. Interpretar um produto, portanto, significa realizar inferências sobre os caminhos utilizados na demonstração daquele desempenho.

Dessa forma, embora a pesquisa aqui proposta verse sobre o processo de realização de testes de língua – afinal, as percepções e impressões que os examinandos têm sobre um determinado aspecto de uma avaliação é uma questão voltada para o processo, para a maneira como esses sujeitos percebem o teste e como elas orientam o modo como o realizam –, de maneira alguma pode-se dizer que se trate de uma pesquisa voltada “totalmente” para o processo. As considerações e constatações feitas visam principalmente a ganhos na compreensão da autenticidade nas avaliações de língua de uma maneira global, e as implicações sugeridas dizem respeito mais diretamente às percepções que orientam o

realizar da avaliação (processo), mas num sentido mais amplo procuram entender também como elas influenciam no resultado geral demonstrado num teste (produto).

Como já indiquei no início da seção, as pesquisas de caráter quantitativo sempre tiveram mais espaço na área de avaliação, embora recentemente o enfoque qualitativo também tenha se firmado como perspectiva importante de trabalho. Algo semelhante também se dá com relação a foco no processo e no produto, sendo o foco no produto mais investigado na área de avaliação de línguas. Um dos motivos para tal é o fato de que o próprio objeto *avaliação* é uma ferramenta de obtenção de produtos de aprendizagem. Ela é usada para se averiguar o desempenho num determinado domínio lingüístico, fruto de algum tipo de aprendizagem (sistemizada ou não) e, assim, a investigação da avaliação a partir do enfoque em seu produto se mostra como um caminho até óbvio.

No entanto, entender o processo de avaliação também é necessário, pois, como já defendi acima, a avaliação não é feita só de processo ou só de produto. Uma crítica que poderia ser feita sobre o foco no processo, por exemplo, é a de que, principalmente com relação a exames externos de proficiência em LE, na verdade o que interessa é apenas o produto da avaliação: uma vez que corretores desses exames têm apenas em mãos as respostas dadas aos itens e é a partir delas que julgam a proficiência de um candidato, não interessa saber quais meios ele utilizou para chegar até ali, pois é somente a partir do *produto* que inferências serão feitas sobre sua capacidade lingüística.

Embora coerente, esse tipo de crítica se mostra um pouco limitada quanto ao possível campo de aplicação dos elementos trazidos por uma pesquisa de processo. Realmente, na correção dos itens, os corretores contam apenas com dados do produto da avaliação, mas a compreensão de outros aspectos do complexo da avaliação, tais como as perspectivas dos examinandos, pode contribuir em outras frentes nas quais a questão da avaliação de línguas está envolvida. Buscar uma compreensão mais global dos instrumentos avaliativos traz conseqüências para o modo como orientamos nossas visões para as avaliações de língua, como as formulamos e corrigimos, para como aplicamos seus resultados e para o valor que elas assumem em diferentes contextos.

Tendo demonstrado como o problema de pesquisa direcionou o viés metodológico, passo agora a descrever o método de coleta de dados (entrevistas semi-

estruturadas), os materiais utilizados (três tipos diferentes de avaliações de leitura em LE) e os sujeitos de pesquisa. É necessário deixar claro que esses direcionamentos metodológicos foram sendo refinados à medida que a pesquisa ia tomando corpo, através das leituras para a definição do referencial teórico, discussões e reflexões individuais e coletivas (sessões de orientação, em grupos de pesquisa, em congressos e seminários) e a realização de uma primeira coleta de dados piloto, onde materiais e métodos foram testados e ajustados para a coleta efetiva. O estudo piloto realizado e suas implicações teóricas e metodológicas serão apresentados no *Capítulo III – A análise e discussão dos dados*. O que descrevo a seguir é o desenho metodológico já imbuído dessas contribuições, que constitui a base da coleta de dados da pesquisa.

1.4.1. Instrumento de coleta de dados: entrevistas semi-estruturadas

A natureza da pesquisa aqui proposta pedia um método de coleta de dados que desse espaço para que opiniões e percepções sobre a autenticidade de avaliações de leitura em LE fossem expressas. Mais que isso, era necessário investigar *se* o examinando já tinha algum tipo de concepção sobre o que seria autenticidade da avaliação, e como esse conceito se configurava. A utilização de entrevistas semi-estruturadas foi o método mais indicado, pois, embora também corra o risco de “contaminar” os dados (como qualquer método) – sendo necessária habilidade para empreender a coleta de maneira a não-influenciar o entrevistado – é um método que permite tentar observar nos sujeitos suas opiniões, as sutilezas na variedade de dados apresentados, aspectos mais particulares de suas percepções e, mais importante, permite incitar o sujeito a discorrer sobre suas percepções de uma maneira mais livre, sem as limitações de um questionário ou *check-list*. Como o objetivo da pesquisa era o de tentar buscar no examinando o que constituía a *sua* perspectiva com relação à autenticidade de uma avaliação de língua, as entrevistas semi-estruturadas conformaram-se como o método mais adequado na busca desse objetivo.

As entrevistas⁴ foram articuladas em duas partes: na primeira, buscando traçar o perfil dos sujeitos, ele discorre mais livremente sobre sua experiência com testes de LE, sobre como geralmente os realiza e suas opiniões sobre diferentes itens avaliativos. A

⁴ Cf. APÊNDICE I.

seguir, eles são incitados a analisar detalhadamente três avaliações de leitura distintas e a relatar suas percepções sobre elas. A decisão de solicitar a análise e pedir para que os sujeitos descrevessem suas percepções sobre os itens sem a exigência de que efetivamente os realizassem é justificada com mais cuidado na seção 1.4.3. *Sujeitos*, e versa sobre a natureza heterogênea das proficiências dos sujeitos e a introdução de variáveis aos dados que poderiam conduzir a considerações não-confiáveis, quando se utiliza diferentes instrumentos avaliativos para cada nível de proficiência abrangido. A escolha desses instrumentos avaliativos, utilizados como material de análise durante a entrevista, é discutida a seguir.

1.4.2. Materiais de coleta de dados: avaliações com diferentes níveis de integração de habilidades

As avaliações de leitura que os sujeitos deveriam analisar, e sobre as quais versaria a segunda parte da entrevista, precisavam apresentar diferentes graus de integração de macro-habilidades, para se investigar como os examinandos percebiam essas características nos itens. Como o objetivo era refletir sobre o conceito de autenticidade presente nessas avaliações, necessária era também a utilização de itens que figurassem alguns “usos reais” da leitura em LE e que se prestassem como avaliação de leitura para que os sujeitos pudessem tecer suas considerações sobre a autenticidade percebida (ou não) nesses instrumentos.

Obviamente, definir o que é um “uso real” da leitura em LE é uma tarefa complicada e que pode se revelar inadequada se generalizações são feitas. Se as hipóteses iniciais de pesquisa versam justamente sobre a relatividade da autenticidade de uso da língua, seria contraditório tentar defini-la em termos absolutos nesse momento. No entanto, temos exemplos inúmeros, em materiais didáticos, exames de proficiência e currículos escolares, sobre o que seriam “avaliações comunicativas”, que a princípio pretendem se aproximar dos usos autênticos de língua. Baseados nos tipos de itens e situações encontrados nesses materiais, pode-se traçar um perfil do que geralmente se toma como reprodução do “uso real de língua”: são itens que procuram refletir situações diversas do cotidiano, tais como estudos, trabalho, viagens, convivência familiar etc. Esses itens muitas

vezes apresentam o contexto da tarefa que propõem, informando seu propósito, quem são os participantes virtuais envolvidos etc.

Os itens⁵ utilizados podem ser assim caracterizados:

- o item identificado por *Tarefa 1*, constituído por uma passagem textual relativamente curta e questões de múltipla escolha, foi retirado de um antigo exame TOEFL, anterior às mudanças que vêm sendo implementadas;

- o item que se propunha como relativamente integrado foi construído como uma proposta de resumo em Língua Materna (doravante LM) que procurasse sintetizar as idéias principais de um texto em LE. O texto escolhido para constituir a *Tarefa 2* foi retirado de um portal conhecido de internet, numa seção de artigos de vulgarização científica retirados da revista *Psychology Today*.

- o item de habilidades integradas, *Tarefa 3*, teve seu texto retirado da mesma fonte do texto da *Tarefa 2*. Assim, como para a *Tarefa 2*, o texto para a *Tarefa 3* deveria ser autêntico, retirado de fontes acessíveis aos sujeitos, e no qual o gênero textual fosse passível de figurar num exame de proficiência⁶.

Tendo recorrido sobre os materiais e métodos que compõem a metodologia de coleta de dados, apresento na seção seguinte uma descrição dos perfis dos sujeitos que participaram da pesquisa.

1.4.3. Sujeitos

Sendo a intenção da pesquisa de cunho qualitativo investigar de maneira mais detalhada um escopo menos amplo que o de pesquisa quantitativa, um número de 8 sujeitos participantes foi considerado adequado aos propósitos do estudo.

A escolha dos sujeitos foi baseada nos direcionamentos do próprio problema de pesquisa e nos pressupostos iniciais sobre sua natureza, como já busquei explicitar nas seções anteriores. No caso dos sujeitos, essas orientações teóricas refletiram-se na escolha

⁵ Cf. APÊNDICE III.

⁶ Na Tarefa 1, texto e itens foram retirados de um exame TOEFL; assim, diferentemente das Tarefas 2 e 3, não houve a seleção do texto e posteriormente criação dos itens especificamente para a pesquisa, pois texto e itens foram selecionados conjuntamente de um exame já pronto. Por isso, a descrição dos critérios para seleção do texto feita acima (textos autênticos, passíveis de serem encontrados num exame etc.) aplicam-se somente para as Tarefas 2 e 3.

de perfis que representassem diferentes necessidades e vivências com relação à leitura em LE e a avaliações de LE. Tomando a autenticidade das avaliações como relativa, dependente do tipo de contexto e situação de uso de língua a que estamos nos referindo, natural foi buscar contemplar diferentes tipos de perfis de sujeitos, que poderiam expressar diferentes percepções sobre a autenticidade.

Abarcar sujeitos de variados perfis é um dos motivos pelos quais adotou-se a análise dos itens, e não sua realização efetiva. Diferentes níveis de proficiência exigiriam itens adequados às diferentes faixas de proficiência abarcadas. Solicitar que sujeitos de diferentes proficiências realizassem uma mesma *Tarefa 1*, por exemplo, poderia enviesar os dados: buscando adequá-la aos níveis mais básicos, níveis mais avançados poderiam considerá-la fácil demais, alterando a percepção que poderiam ter sobre o item com relação ao seu formato, sua integração de habilidades e seu nível de autenticidade. O mesmo poderia se dar no caminho inverso, adaptando-se o item aos níveis mais avançados.

No mesmo sentido, utilizar mais de um tipo de item para cada grau de integração de habilidades – mais de uma Tarefa 1, por exemplo, usando o mesmo formato de múltipla escolha, mas adequando a exigência do item à proficiência lingüística dos sujeitos – introduziria variáveis que poderiam enviesar os dados. Uma avaliação deve ser adequada ao nível de proficiência esperado do aluno/sujeito/candidato: textos e itens devem estar coordenados com o que se espera que essa proficiência permita a ele realizar com a LE. Dessa forma, se sujeitos com diferentes proficiências fossem realizar os itens, obrigatório seria a utilização de diferentes textos e itens de avaliação, adequados aos diferentes níveis de proficiência. Esse procedimento ampliaria o número de variáveis presentes, dada a utilização de diferentes instrumentos e a falta de garantia que eles se equivalessem em termos de nível de exigência, grau de dificuldade dos textos e proximidade com usos de língua feitos na vida real.

Para que pudéssemos ter um parâmetro comparativo do nível de proficiência dos sujeitos, um teste cloze⁷ foi aplicado após as entrevistas. Testes cloze consistem num texto em que são apagadas algumas palavras, e cuja tarefa do examinando é completar essas lacunas apropriadamente. Os testes cloze foram alvo de grande discussão na área de

⁷ Cf. APÊNDICE II.

leitura e avaliação por alguns anos, tomados por alguns como uma medida estreita de avaliação. Estudos posteriores (Oller, 1979, Alderson 1979, Shohamy 1982), no entanto, reconheceram o teste cloze como um instrumento avaliativo abrangente. Embora fornecendo uma avaliação relativamente grosseira de proficiência, não se prestando a detalhes e sutilezas sobre a proficiência dos examinandos que outros instrumentos avaliativos contemplam, o cloze já foi usado com eficiência como teste de proficiência em vários estudos, como o de Scaramucci (1995).

O teste cloze aqui usado foi o mesmo proposto por Scaramucci (ibid., p. 117), identificado pela autora como *Cloze II*, um instrumento validado em sua pesquisa, e cujo propósito de mensuração da proficiência geral dos sujeitos ia ao encontro do presente estudo. O teste foi elaborado a partir de um texto da revista *Time* e tem 36 palavras apagadas de acordo com um procedimento de seleção racional de apagamento, que apaga palavras segundo seu *valor comunicativo* para o contexto, sua *recuperabilidade* e *relevância* (Scaramucci, ibid, p. 116). Os critérios de correção também são os mesmos adotados pela autora, descritos de forma detalhada no Capítulo III.

Outras informações, além do nível de proficiência, permitem traçar os perfis dos sujeitos participantes, a saber: tipo de formação acadêmica, tipo de atividade que exercem, que necessidades têm com relação à LE, usos de leitura em LE que fazem no dia-a-dia, suas experiências com a realização de avaliações em LE em geral e de leitura em LE mais especificamente, assim como os tipos de aprendizagem que vivenciaram em LE. Descrições detalhadas de como essas características se configuram em cada um dos sujeitos são feitas no *Capítulo III – Análise e discussão dos dados*.

1.5. Estrutura da dissertação

Neste primeiro capítulo, é traçado um esboço geral da dissertação, com a contextualização e justificativa do problema de pesquisa, a questão que orienta o trabalho e é descrita a metodologia empregada na sua realização.

No Capítulo II, apresento o embasamento teórico do trabalho, discorrendo sobre os conceitos de integração de habilidades, autenticidade na avaliação e percepções

sobre a avaliação, discutindo alguns estudos já feitos sobre esses tópicos que auxiliam na construção da presente pesquisa.

No Capítulo III, descrevo a coleta de dados, apresento a análise deles e a discussão dos resultados.

Por fim, no Capítulo IV, vêm as considerações finais, implicações e limitações da pesquisa, e possíveis caminhos futuros para o prosseguimento de estudos na área.

CAPÍTULO II

FUNDAMENTAÇÃO TEÓRICA

No capítulo anterior, busquei situar o problema de pesquisa dentro de uma perspectiva mais ampla, ressaltando a importância que a avaliação tem no contexto social como ferramenta que atua na tomada de decisões em diversas esferas e que, portanto, influi diretamente nos rumos tomados por pessoas e instituições. Dentro desse contexto, indiquei os questionamentos que orientam a pesquisa e descrevi a metodologia com a qual se busca respondê-los.

Neste capítulo, discorro sobre os pontos que considero essenciais na fundamentação teórica deste trabalho, já indicados no fim do capítulo anterior, a saber: o conceito de integração de habilidades, a autenticidade na avaliação e a perspectiva dos examinandos. Imbricados nesses três pilares, apresento as discussões sobre visões de língua, leitura e avaliação, assim como conceitos próprios do campo da avaliação de línguas, tais como *validade* e *efeito retroativo*. Procuo discorrer sobre os mais importantes para a construção desse arcabouço teórico e relacioná-los ao problema de pesquisa, mostrando como eles atuam em sua fundamentação. Início esse percurso a partir da reflexão sobre alguns aspectos relativos ao estudo e à prática da avaliação de leitura em LE já abordados em estudos anteriores.

Estudos realizados em países com grande tradição de pesquisa na área de avaliação de línguas, como Inglaterra, Estados Unidos e Austrália, já se debruçaram sobre a investigação de diversos aspectos da avaliação de leitura (cf. Alderson 2000). No Brasil, no entanto, estudos que busquem refletir sobre a questão levando em consideração os fatores que a particularizem são poucos. Essa é uma lacuna que precisa ser preenchida, visto que a avaliação de leitura em LE ocupa não apenas um lugar importante nos contextos escolares, mas é o principal instrumento avaliativo e seletivo em diversas esferas sociais.

Este trabalho versa sobre o conceito de avaliação autêntica segundo a perspectiva das pessoas que se submetem a diferentes tipos de testes e exames de leitura

em LE. Entretanto, logo de início, é necessário admitir que esse é um conceito controverso, visto que os usos da leitura em LE são tantos e tão variados que seria impossível eleger uma “avaliação autêntica” por definição, válida para todo e qualquer contexto. Uma discussão mais detalhada do conceito é feita mais adiante, mas é necessário ressaltar já aqui que, apesar dos testes de leitura em geral muitas vezes não refletirem os diferentes objetivos da leitura realizada na vida cotidiana, eles constituem um elemento importante em diversos meios sociais.

A leitura na “vida real” – consultar a lista telefônica, ler revistas, romances, jornais, aprender através de uma receita escrita um novo prato culinário, etc – é motivada pelos mais variados propósitos, não para ser avaliada como um fim em si mesma. A própria situação de avaliação de leitura já constituiria um situação não-autêntica. Embora seja essencial reconhecer que a avaliação, por sua própria natureza, seja essencialmente diferente das outras “situações reais de leitura”, uma vez que cenário, participantes, tempo, recursos, objetivos etc. são diferentes para cada contexto, sua importância reside em ela própria constituir uma dessas “situações reais”. As pessoas, além de toda leitura extensiva (*extensive reading*) que realizam, também precisam ser avaliadas para ingressar numa universidade, arrumar empregos, serem promovidas de cargo etc. Como afirma Alderson (2000, p. 27), (...) *a sociedade precisa de avaliações. A avaliação é um evento diferente de muitas outras atividades, mas é valorizada pela sociedade por seus próprios méritos, e é por isso justificada*⁸.

Nesse sentido, investigar questões relativas à avaliação de línguas e o quanto ela se aproxima (ou não) das situações cotidianas de uso da leitura em LE tem importância para o avanço das pesquisas na área de avaliação — ainda permeada por muitas interrogações e necessidades de estudos aprofundados — e também para a adequada aplicação desse recurso, contribuindo para a reflexão sobre as implicações de sua utilização em áreas como ensino/aprendizagem de línguas e como instrumento de seleção e ingresso em campos educacionais e profissionais.

⁸ Citação original: “(...) *society needs assessment. Assessment is different from many other activities, but it is valued by society for its own sake, and is therefore justified.*”

A discussão sobre a semelhança que as avaliações de leitura possam ter com o modo como a leitura é realizada no dia-a-dia – ou, de maneira mais ampla, a semelhança que as avaliações de LE possam ter com o modo como a LE é efetivamente usada – leva a pensar sobre como esses construtos são postos em prática na “vida real” e nas avaliações. Como já afirmei, e discuto nas seções seguintes, a questão da autenticidade da avaliação está relacionada a isso, e a da integração ou isolamento de habilidades, ou macro-habilidades, também se coloca como fator marcante na constituição da correspondência entre “vida real” e “avaliação”. Assim, e os itens avaliativos vistos segundo essa ótica são brevemente discutidos na próxima seção.

2.1. Integração e isolamento de habilidades

Para discutir o conceito de isolamento e integração de habilidades, em primeiro lugar, na subseção *2.1.1 Surgimento e evolução do conceito de integração e isolamento de habilidades*, busco delinear uma perspectiva histórica do surgimento e evolução da idéia de integração de habilidades na área de avaliação de línguas, tomando por base estudos que se dedicaram à discussão desses termos. Em seguida, na subseção *2.1.2. A atual discussão acerca da integração e isolamento de habilidades*, como o título sugere, tomo para reflexão o que propõem pesquisas recentes, buscando pensar sobre como o conceito é visto e usado hoje em dia pelos teóricos da área. Finalizando a seção 2.1, em *2.1.3. Integração de habilidades nos exames de proficiência*, procuro ilustrar e contextualizar as questões discutidas nas subseções anteriores através de exemplos do lidar com a idéia de integração de habilidades encontrados em exames de proficiência. Nessa subseção, também, indico as amarras teóricas que ligam o conceito de integração de habilidades ao conceito de autenticidade, que é discutido em seguida na seção *2.2. Autenticidade na avaliação*.

2.1.1 Surgimento e evolução do conceito de integração e isolamento de habilidades

Carroll (1961) foi o primeiro estudioso da área a referir-se ao termo *teste integrativo* (*integrative test*), quando discutindo a então novidade das avaliações comunicativas (*communicative testing*). A idéia de integração de habilidades nasceu ligada

à idéia de avaliação comunicativa, procurando estabelecer um contraponto às avaliações de base mais estruturalista que vigoravam até então, cujo principal método de testagem era o isolamento de elementos lingüísticos, avaliados à parte do contexto onde eram produzidos ou recebidos nas comunidades discursivas. Os estudiosos da área passaram a denominar esse tipo de avaliação isolada como *discrete point tests*, em oposição à avaliação não-isolada, que procurava integrar habilidades, os *integrative tests*.

Em português denominados *isolados*, os *discrete point tests* caracterizam-se pela avaliação isolada de componentes do sistema lingüístico, retirando-os de seu contexto “natural”. Itens isolados de avaliação podem se concentrar em aspectos fonológicos, gramaticais ou lexicais de uma língua, porém sempre de uma maneira muito específica, não levando em consideração outros aspectos que fazem parte do contexto em que esses elementos existem. Por exemplo, suponha-se que se pretenda avaliar o conhecimento dos examinandos sobre a ortografia de palavras que tenham S ou SS. Um possível teste isolado seria apresentar uma lista dessas palavras em que fossem apagados os S ou SS de suas grafias (por exemplo: “pá__aro”, “ca__amento” e “a__oalho”) e solicitar aos examinandos que completassem as lacunas com o uso correto de S ou SS.

Uma das principais considerações que podem ser feitas sobre um item como esse é, como já afirmei acima, a falta de contextualização da avaliação. A preocupação não é exatamente com o *uso* que se faz habitualmente do conhecimento da grafia das palavras – para se escrever textos de diversos tipos, com diversos propósitos, para diversos interlocutores etc. –, mas a testagem específica do conhecimento de um tipo de elemento lingüístico separado dos contextos em que é produzido e usado.

O exemplo apresentado acima seria um dos tipos de itens isolados em sua versão mais extrema. Como oposição a esse tipo de avaliação foi que as avaliações comunicativas começaram a ganhar força, e daí surgiu a idéia de avaliações que não isolassem elementos de seus contextos, mas que os apresentassem integrados aos outros aspectos que compõem o modo habitual em que esses elementos são produzidos. Surgia, então, o conceito de *itens integradores*, no sentido de integração de mais de um desses elementos na avaliação.

O conceito de itens integradores evoluiu muito desde então e, hoje, é necessário esclarecer bem o tipo de “integração” a que estamos nos referindo quando trabalhando com esse tópico, dadas as variadas nuances de integração existentes nas avaliações.

A princípio, qualquer item que não se concentre em um único elemento lingüístico poderia ser considerado integrador. Essa definição, no entanto, não parecia ser suficiente para dar conta de um campo tão complexo quanto o das avaliações comunicativas. Oller (1979), buscando caracterizar melhor o tipo de itens que se conformavam a essa nova visão de avaliação, estabeleceu diferenças entre os termos *itens isolados*, *itens integradores* e *itens pragmáticos*. As definições do autor sobre como se caracterizariam cada um desses tipos de itens traça as diferenças básicas entre eles, embora a sobreposição parcial que existe na definição de itens integradores e itens pragmáticos não seja deixada de lado.

O termo *testes pragmáticos*, usado por Oller para definir um tipo de item que procurava não apenas integrar elementos, aspectos ou habilidades, mas fazia isso tentando seguir uma lógica própria do uso real de língua, procurava conferir um caráter “mais comunicativo” aos itens avaliativos. O período “sociolingüístico” da avaliação (Shohamy 1985a, p.4) caracterizou-se como o momento em que gradualmente os ideais afastavam-se dos testes isolados de outrora, e a definição de testes pragmáticos constitui uma dessas buscas do item cada vez menos isolado, mais atento a fatores do contexto. Percebe-se que, de modo análogo, essa transição revela as mudanças que ocorriam nas visões de linguagem não apenas na área da avaliação de línguas, mas nos estudos da linguagem em geral, que procuravam direcionar seu olhar para a língua de uma maneira mais comunicativa e menos estruturalista.

Oller (1979) define testes pragmáticos como aqueles em que a avaliação integra elementos do sistema lingüístico de uma maneira que se conforma aos padrões em que a língua é usada genuinamente. Assim, para ele, testes cloze e ditados seriam exemplos de testes pragmáticos, pois não tomam elementos lingüísticos isolados na avaliação – ao contrário de testes isolados, em que se avalia, desprovido de contexto, o conhecimento do examinando sobre passado dos verbos irregulares, por exemplo. Num teste cloze, uma das palavras apagadas poderia ser um verbo irregular no passado – e, assim como no teste

isolado, o cloze avaliaria o conhecimento do passado de verbos irregulares. Isso se daria, porém, de uma forma contextualizada, pois esse elemento estaria inserido num texto, carregado de significado, sendo usado como normalmente as pessoas o usam.

Outro fator que poderia ser apontado para fortalecer o caráter integrador e pragmático do cloze, por exemplo, seria a integração que ele promove na avaliação de diferentes tipos de conhecimentos lingüísticos. Existem diversos tipos de cloze, com diferentes critérios de apagamento de palavras e de correção (ver Scaramucci 1995, p.115). No entanto, algo em comum que eles teriam, além do formato, seria o fato de o cloze avaliar a proficiência em leitura em mais de um nível⁹, pois dependendo da palavra apagada, exigiria a mobilização de um tipo de conhecimento – como, por exemplo, conhecimento gramatical, lexical, ortográfico etc. Assim, um mesmo item avaliativo integraria a avaliação de diversos conhecimentos e capacidades lingüísticas.

Embora estudos já tenham reconhecido a abrangência dos testes cloze (cf. Alderson 1979, Shohamy 1982) não apenas como um instrumento de avaliação de leitura, mas mesmo de proficiência geral – dado o alto índice de correlação com outros tipos de teste que o cloze demonstra ter, quando observados os devidos rigores de preparação e aplicação –, as novas nuances que o termo “integrador” assumiu fez com que sua ampla aceitação como instrumento avaliativo fosse questionada. Também os ditados, assim como outros tipos de itens citados por Oller (op. cit.) como “pragmáticos” não parecem corresponder à noção de integração e de avaliação comunicativa que hoje é comum. Estudos mais recentes já se referem à integração como envolvendo o componente pragmático em si, sem a necessidade de se referir a “testes pragmáticos” como algo “a mais” que integrador. Além disso, itens como o cloze e ditados não correspondem exatamente ao ideal de avaliação que a busca por testes que correspondessem cada vez mais aos ideais comunicativos de avaliação almejava – isto é, itens que buscassem contemplar aspectos sociolingüísticos, cujo formato das tarefas se assemelhasse ao modo como a língua

⁹ Embora essa seja uma afirmação controversa, alguns teóricos defendem a idéia de que alguns apagamentos são recuperados através da compreensão no nível da sentença, enquanto outros exigem compreensão mais abrangente, do texto como um todo; de qualquer forma, não é possível afirmar isso sobre qualquer cloze, pois como já afirmei, existem vários tipos de cloze, com textos mais fáceis ou mais difíceis, apagamentos aleatórios ou selecionados etc.

é usada no dia-a-dia, que fornecessem amostras de desempenho que se aproximassem do tipo de desempenho esperado nas interações autênticas.

Embora hoje o termo integrador represente mais do que o que Oller definia no final da década de 1970 como integrador, podemos dizer que sua definição representou uma das primeiras tentativas da já mencionada busca por avaliações mais condizentes com os usos reais de língua. Morrow (1979) também já aponta a necessidade de situar as avaliações dentro de um contexto, integrando insumo autêntico a tarefas que simulassem os tipos de tarefas que os examinandos deveriam desempenhar nos contextos exteriores ao teste. Podemos dizer que essa foi a principal vertente pela qual seguiu a evolução do termo “integrador” no campo da avaliação de línguas.

Shohamy (1985a p.4) denomina o período pós-item isolado como “Sociolingüístico-Integrador” e aponta a busca cada vez maior pela integração de habilidades, em diversos níveis, como uma de suas principais características:

*“A avaliação está deixando para trás a ênfase em habilidades lingüísticas separadas, em direção à comunicação prática, que representa a combinação de todas as habilidades. Testes incluem tarefas comunicativas e a avaliação de fatores culturais e sociolingüísticos. As avaliações incluem tarefas reais de comunicação. O aprendiz de línguas deve ser sensível a quem diz o que, para quem, em que momento e de que maneira. A avaliação inclui também adequação, registro e capacidades comunicativas.”*¹⁰

Como se vê, esse é um sentido muito mais completo de integração. Nota-se também, imbricado nele, a noção de autenticidade, em diversos níveis: na atenção ao formato das tarefas, que deveria se assemelhar aos da vida real; atenção à comunidade discursiva onde são produzidos e recebidos os enunciados do insumo lingüístico usado no teste – retirado, de preferência, de fontes autênticas, ou seja, não criado especialmente para a avaliação; na importância atribuída à capacidade de adequação lingüístico-social e

¹⁰ Citação original: “Testing is moving away from emphasizing separate linguistic skills toward practical communication which represents a combination of all skills. Tests include communicative tasks and testing of cultural and sociolinguistic factors. Testing includes real tasks of communication. The language learner should be sensitive to who says what, to whom, at what time and in what fashion. Testing includes also appropriateness, register, and communicative abilities.” (Shohamy 1985a, p. 4)

também a outras capacidades comunicativas, tais como competências estratégicas ou sociolingüísticas, por exemplo.

Um dos sentidos que podemos depreender da citação de Shohamy é o de integração de habilidades como avaliação de macro-habilidades de forma integrada – testes que integrem leitura e escrita, por exemplo, ou leitura e produção oral, dentre outros. Outro é o sentido mais amplo, em que são avaliados conhecimentos do examinando sobre o tipo de texto a ser lido, o tipo de gênero a ser produzido na tarefa, a adequação aos propósitos e interlocutores sugeridos etc. Uma tarefa que solicite a leitura de um artigo de jornal e, baseado nisso, a escrita de uma carta ao editor, por exemplo, integra as macro-habilidades de leitura e escrita, testando compreensão do texto e produção escrita e o domínio da linguagem que o examinando tem, como conhecimento de estruturas sintáticas, léxico etc.; mas, mais do que isso, avalia o conhecimento do examinando sobre os gêneros dos textos – “artigo de jornal” e “carta” –, sobre o registro a ser usado dado o interlocutor sugerido – “editor” –, sobre os propósitos e limites contextuais impostos nesse tipo de situação e o que eles significam em termos de uso de língua.

Alderson (2000, pp.25-27), já trabalhando com o termo “integração” nessa perspectiva mais recente, afirma que a visão de leitura como socialmente imbricada com outras habilidades, numa perspectiva de letramento como prática social, traz implicações para o modo como a leitura é avaliada. Essa nova postura, conseqüentemente, questiona o modo “isolado” com que alguns testes avaliam a leitura. O autor apresenta alguns dos argumentos que alguns teóricos sustentam para defender o isolamento ou a integração de habilidades, procurando refletir sobre as diferentes perspectivas da questão. A favor do isolamento, estariam vantagens como evitar a “contaminação¹¹” de habilidades, a avaliação enviesada pela provisão de insumo escrito (que facilitaria o processo de escrita em LE e até estimularia a simples cópia); ela evitaria também a complexidade que é própria da integração de habilidades e que poderia resultar em testes mais “difíceis” devido à construção do item, e não pela “dificuldade” inerente ao construto avaliado. Já a integração seria mais condizente com uma visão comunicativa de linguagem, trabalha a leitura dentro

¹¹ O conceito de contaminação na avaliação (*muddied measurement*, Weir 1990) é discutido nas páginas 44 e 45.

de um contexto letrado mais amplo (e, por isso, mais significativo), não vê a leitura como simples habilidade cognitiva, mas também como prática social.

Percebe-se que o autor defende uma visão de “avaliação integrada” não somente como integração de macro-habilidades, mas também como integração da leitura a práticas letradas diversas, a um contexto de uso que extrapola o lingüístico e entra no universo social mais amplo, no qual a leitura cumpre diferentes papéis que devem ser levados em consideração na construção de avaliações. E, embora concorde que não há como considerar a leitura somente como uma capacidade cognitiva, sugere que maior investigação é necessária na área para que se possa ter uma idéia mais clara dos resultados práticos que o isolamento ou a integração têm quando empregados em avaliações de leitura em LE, pois embora diferentes níveis de integração sejam efetivamente usados como parte dos mais diversos tipos de avaliação de língua existentes em meios escolares e não escolares, ainda existe muita divergência sobre os processos envolvidos na avaliação através desses itens (cf. Alderson 2000, p. 26). A validade, por exemplo, é um dos aspectos envolvidos nessa questão.

A *validade* de um item avaliativo pode ser definida como a medida em que esse item fornece uma amostra significativa do domínio da LM ou LE que se propõe a avaliar. Pretendendo-se avaliar a capacidade de leitura numa LE, deve-se assegurar que essa habilidade está contemplada adequadamente na avaliação – um item que se concentrasse sobre a averiguação do conhecimento de itens isolados de gramática presentes num texto, por exemplo, teria validade contestável como avaliação de leitura. Na literatura de avaliação, as nomenclaturas *validade de conteúdo*, *concorrente*, *preditiva*, entre outras, são utilizadas para diferentes tipos de validade¹². Como já afirmei anteriormente, aspectos como concepções de linguagem, de leitura e de avaliação estão imbricados nessa questão, e devem ser tomados como fatores de uma reflexão que busque investigar o uso desses diferentes tipos de itens.

¹² Para uma discussão mais ampla do conceito, ver Hughes (1990) e Shohamy (1985a, p. 74-75).

2.1.2. A atual discussão acerca da integração e isolamento de habilidades

Os usos dos diferentes tipos de itens, como discuti na subseção anterior, são um ponto controverso na literatura da área. Uma maior integração de habilidades, por exemplo, é advogada por alguns estudiosos como sendo mais coerente com uma visão de linguagem como um meio para agir no mundo¹³. As pessoas lêem para realizar as mais variadas atividades – para fazer uma receita culinária, para aprenderem algum conceito, para responder uma carta ou e-mail, para se prepararem para falar em público etc (cf. Clark, H.H. 2000). Para autores como Hill & Parry (1992), o uso da língua não se dá na mobilização de macro-habilidades separadas, mas é a simultaneidade do ler, escrever, falar e compreender a língua que constituem esse uso e, portanto, uma avaliação que as integre é muito mais “real” do que uma que tente avaliá-las separadamente.

Itens de abordagem mais isolada, por outro lado, também têm seus defensores. Um dos fatores mencionados que os tornariam mais apropriados como instrumentos avaliativos seria a “contaminação” que itens integradores acabam causando. Um item em que apenas uma habilidade é avaliada – somente leitura ou escrita, por exemplo – teria como principal vantagem o fato de ser um meio mais “puro” de acesso à capacidade de desempenho do examinando numa determinada habilidade. Examinar a compreensão em leitura através da escrita, por exemplo, poderia revelar-se uma medida enviesada, “contaminada”: a deficiência na produção escrita poderia atrapalhar a avaliação da leitura. Assim, a avaliação não cumpriria seu papel de, corretamente, avaliar, diagnosticar os pontos fortes e fracos de um examinando.

Um outro ponto sobre o qual é necessário refletir quando abordando a questão de integração ou isolamento de habilidades é se a avaliação independente de uma habilidade é realmente indicativa de uma concepção de linguagem idealizada e irreal. Embora, a princípio, a integração de habilidades pareça refletir de maneira mais fiel o modo como a língua é usada, em muitas situações o que realmente presenciamos é o uso isolado de habilidades. Muitas pessoas utilizam uma LE em suas vidas apenas para ler, por exemplo, não demonstrando a necessidade de escrever ou falar com alguém em outra língua

¹³ O exame CELPE-Bras é um dos que tem esse tipo de princípio norteador. Para uma discussão mais aprofundada, ver Scaramucci 2004b

que não a LM. É necessário estar atento a esse tipo de fator quando debatendo a integração de habilidades, para não se correr o risco de fazer afirmações generalizantes e categóricas, sem a devida atenção a sutilezas e variações de uso de contexto para contexto.

Shohamy (1985a, p.5), dentre outros, salienta um movimento de tentativa do abandono de abordagens mais tradicionais em direção a abordagens mais “sociolingüísticas”. Desde então, a busca pelo que viria a ser essa abordagem sociolingüística na avaliação só fez crescer, impulsionada por estudos na mesma direção em outras subáreas da Lingüística Aplicada. O impasse, porém, sobre os aspectos assinalados acima, que cercam a definição e o uso das avaliações de leitura, tornam essas avaliações, muitas vezes, objeto de uso equivocado e/ou não satisfatório. Compreender os fatores e processos que estão envolvidos numa avaliação pode, portanto, trazer ganhos para a maneira como ela é formulada, aplicada e corrigida. E isso, por sua vez, pode influenciar os efeitos retroativos que possam causar na sociedade – tanto no que se refere ao ensino quanto no que se refere aos impactos mais imediatos, como o modo como o resultado de um exame interfere no funcionamento de instituições sociais e na vida das pessoas mais diretamente (cf. Scaramucci 2004a).

Como se vê, não apenas o conceito do que é a integração e o isolamento de habilidades é complexo, mas também o são as extensões e implicações que as práticas desses diferentes métodos avaliativos produzem. Na busca de maior clareza nesse sentido, Lewkowicz (1997), num texto que procurava caracterizar o estado da arte da integração de habilidades na avaliação, faz uma diferença terminológica entre *integrativos* e *integrados* (*integrative* e *integrated*), propondo os termos como as extremidades opostas de um continuum. De um lado, estariam os testes *integrativos*, que seriam os itens em que há uma certa integração de habilidades e que representam uma medida holística de mensuração da proficiência. Como exemplos, ela cita o que Oller (1979) definiu como testes pragmáticos, o cloze e o ditado.

Já no outro extremo do continuum, os testes *integrados* representariam o máximo de integração de habilidades segundo a definição de integração mais recente, sobre a qual discorri acima, em que não apenas habilidades são avaliadas em conjunto, mas integrados estão também aspectos sociais da linguagem. Nos itens *integrados*, há a

integração de habilidades, de temas e de tipos de capacidades e conhecimentos e lingüísticos, e o insumo proposto serve de base para a produção lingüística do examinando. A avaliação é tomada como um instrumento em que o desempenho demonstrado no teste deve refletir ao máximo o tipo de desempenho esperado em situações reais de uso da língua.

Poderíamos dizer, a partir disso, que o termo *integrativo* se refere àquela definição de integração mais antiga, dada por Oller (1979), que inspirou os primeiros passos em direção aos ideais de avaliações comunicativas do período denominado por Shohamy (1985a, p. 4) como “sociolingüístico-integrador”. À medida que avançamos nesse continuum, encontramos os acréscimos feitos ao longo desse período ao conceito de integração: estariam aí localizados os testes nos quais já existe maior atenção devotada ao formato (que simulasse o formato de uma tarefa real), ao uso de insumo autêntico, à avaliação de aspectos sociolingüísticos e culturais, que procurassem refletir interações e situações autênticas. Essas características estariam presentes num grau menos acentuado nas tarefas que se localizassem próximas ao extremo *integrativo* do continuum, e em grau mais alto nas tarefas próximas ao extremo *integrado*.

Graficamente, poderíamos representar esse continuum pela Figura 1 abaixo, em que os extremos *Integrativo* e *Integrado* seriam representações virtuais de graus de integração de habilidades de um item avaliativo, em seus níveis mais extremos. Dentro desse continuum, situariam-se diferentes níveis de integração, e o Item A configuraria uma tarefa de menor integração de habilidades – mais próxima do extremo *Integrativo* – e o Item B, uma tarefa de maior integração de habilidades - mais próxima do extremo *Integrado*.

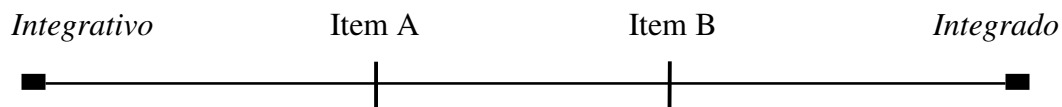


Figura 1: Continuum dos níveis de integração de uma tarefa

2.1.3. Integração de habilidades nos exames de proficiência

Um exemplo prático da integração e seus efeitos pode ser encontrado quando consideramos o exame IELTS (*International English Language Testing System*), de responsabilidade do UCLES (*University of Cambridge Local Examinations Syndicate*) e as mudanças empreendidas em seu formato desde sua implantação como exame de proficiência em inglês para candidatos a programas de estudos e trabalho em países falantes de língua inglesa. O exame continha uma grande integração entre os componentes de leitura e escrita em seu formato até 1995, quando essa relação foi extinta. Um dos pontos advogados em defesa dessa integração no exame era a provisão de insumo escrito que visava a garantir que candidatos de contextos culturais diferentes de onde o exame é preparado não tivessem suas capacidades lingüísticas subestimadas devido à falta de conhecimento prévio nos assuntos sobre os quais deveriam produzir no componente escrito do exame. Esse é um dos motivos dados por Hill & Parry (1992), por exemplo, para defender a integração de habilidades. Os autores discutem a limitação das avaliações em leitura baseadas num modelo “autônomo” de letramento (que assume texto, leitor e a atividade de leitura como unidades “autônomas” e desprovidas de contexto), oposto a um modelo “pragmático” (que vê um complexo enredamento social entre essas três instâncias, além da atribuição de identidades sociais a leitores e autores e contextualização dos ambientes de produção e recepção de textos em comunidades letradas). De acordo com eles, o efeito maléfico de testes baseados num “modelo autônomo” é ainda mais acentuado quando os candidatos submetidos à avaliação dependem de educação formal para o desenvolvimento de suas habilidades lingüísticas e letradas, como é o caso de uma LE.

No entanto, de acordo com os responsáveis pelo IELTS, a validade de construto parecia comprometida com esse formato integrado. A maneira como os candidatos usavam o insumo escrito variava muito e, na maioria das vezes, não era proveitosa, e esses foram argumentos usados para justificar a retirada dessa integração entre leitura e escrita. Segundo eles, não era claro para muitos como esse insumo deveria ser tratado. Dessa forma, muitos consideravam a tarefa como avaliativa de sua habilidade de leitura apenas, não demonstrando adequadamente sua habilidade de escrita; outros realizavam as tarefas

usando apenas seus próprios conhecimentos, sem usar o texto indicado para leitura (Charge & Taylor 1997).

Essa mudança na estrutura do exame foi criticada, embora os que advoguem seus benefícios afirmem que monitoramento cuidadoso durante a fase de pilotagem dos exames tenha demonstrado que os candidatos não são prejudicados com o novo formato (cf. Charge & Taylor 1997, p. 376). As críticas, porém, sugerem que uma das características essenciais da integração de habilidades no exame – a garantia de que elementos de conhecimento temático fossem fornecidos, de modo a não mascarar a habilidade de escrita dos candidatos que não dispusessem deles – foi totalmente perdida com as alterações feitas e que isso é altamente prejudicial para muitos candidatos. Wallace (1997, p.372) sugere ainda que o formato integrado do exame, se tomado para análise pelo aspecto do componente cognitivo da leitura, também contemplava perfeitamente habilidades requeridas nesse processo, como capacidade de realização de síntese, paráfrase ou resumo. Assim, mudanças eram necessárias na grade de correção do exame, e não em seu formato; ela deveria ser reformulada de maneira a valorizar a presença dessas micro-habilidades e penalizar simples cópia, por exemplo. Lewkowicz (1997, pp. 125-126) afirma que, nas primeiras aplicações do IELTS, este era um exame verdadeiramente integrado, mas a desaprovação de muitos professores em relação à quantidade de leitura exigida “abrandou” a integração do exame, e a ligação entre leitura e escrita que passou a existir até 1995 tornou-se apenas aparente. A autora aponta que foi essa integração que se revelou problemática, levando a casos de, por exemplo, candidatos realizarem a tarefa escrita sem precisar se referir ao texto lido – problemas que foram indicados como a causa do completo desligamento entre os módulos de escrita e de leitura nas versões mais recentes do exame.

O caso do exame IELTS ilustra bem alguns pontos positivos e negativos da avaliação integrada de habilidades. Se, por um lado, ela busca assemelhar-se ao modo como a língua é usada no cotidiano e prever de maneira mais direta o desempenho futuro do candidato usando a LE, por outro, ela não é tão transparente quanto a avaliação isolada de uma habilidade em termos do construto mensurado, tanto pelo problema reportado no caso do IELTS quanto pelo que Weir (1990) chamou de *avaliação contaminada (muddied measurement)*. A “contaminação” ocorre quando a avaliação de uma macro-habilidade

sofre interferência de outra, de modo que o produto dessa avaliação como um todo possa demonstrar um desempenho aquém da proficiência usual do examinando. Assim, um examinando bom leitor¹⁴ poderia ser prejudicado por deficiências na sua habilidade de escrita, quando essas duas habilidades estivessem sendo avaliadas de forma integrada, por exemplo; ou alguém com mais dificuldades na compreensão oral poderia ter sua habilidade de escrita subestimada, quando dependesse da compreensão de um trecho de uma conversa para elaborar um texto escrito. No caso de avaliar leitura e escrita conjuntamente, soma-se uma variável a mais a esse problema: o das capacidades de produção e recepção habitualmente não seguirem o mesmo compasso de desenvolvimento, principalmente nos níveis de proficiência mais básicos. Lee (1986), numa pesquisa sobre os efeitos do conhecimento prévio na leitura, afirma que a natureza do método empregado para a coleta de dados – respostas dadas em LM e não em LE num teste de leitura – foi essencial para explicar a diferença entre seus resultados e resultados de pesquisas anteriores, uma vez que geralmente a capacidade de compreender uma LE é maior que a capacidade de produzi-la, e que, por isso, o modo como a compreensão da leitura é avaliada influencia a avaliação dessa compreensão.

Um dos problemas da contaminação da avaliação mais evidente é a perda da função diagnóstica da avaliação. Quando ocorre a contaminação, não é possível verificar onde estão localizados os pontos fracos e fortes do examinando. Numa avaliação de sala de aula, por exemplo, isso é primordial, se pensamos na avaliação como um instrumento formativo, de auxílio no estabelecimento de um plano de ação que vise à melhora do rendimento do aluno. Num exame de proficiência, esse efeito é minimizado se o objetivo do teste for estabelecer um nível de proficiência de maneira global.

Ressaltamos aqui que o termo *proficiência* é aqui usado como relativo à situação de uso, e não como um “nível absoluto” que deva ser atingido para que se possa ser considerado capaz de usar a LE. Um indivíduo pode ter proficiência básica,

¹⁴ Embora aqui usado, o termo *bom leitor* não parece muito apropriado, pois pode dar a impressão que um juízo de valor subjetivo é emitido. E também porque, assim como Scaramucci (1990, p.78), *julgamos problemática a caracterização do bom leitor*, pois essa caracterização não pode ser absoluta, mas relativa aos objetivos de leitura, tipo de texto etc. Assim como para a autora, o termo aqui se refere aos sujeitos que se sobressaem nos testes aplicados.

intermediária, avançada, proficiência para ler manuais, para ser comissário de bordo, para ser tradutor – daí nos referirmos a *graus* ou *níveis de proficiência*.¹⁵ Se o teste almejar reportar separadamente o desempenho do candidato em “faixas” de proficiência em cada uma das habilidades, um fator de dificuldade a mais é adicionado, visto que, na integração, as habilidades se mesclam na avaliação, e a contaminação, se não impedisse, complicaria esse processo. Portanto, a adoção de um determinado grau de integração na elaboração de um teste deve estar atenta aos propósitos do mesmo como instrumento avaliativo.

Lewkowicz (1997) discute brevemente os efeitos da contaminação de habilidades demonstrado por Weir (1990), assim como identifica outras questões ainda presentes nesse campo. Entre elas, aponta questões de praticidade do teste, característica que nem sempre é possível de ser alcançada, dada a complexidade e o dispêndio de recursos (humanos, financeiros, materiais) necessários na formulação e aplicação de testes integrados. Sempre apontando a busca de autenticidade lingüística como inerente a avaliações que se propõem integradoras, a autora ainda indica a dificuldade de definir o que realmente é “autêntico” em termos de uso de língua, assim como as questões que daí derivam, como o quão complexo torna-se garantir a confiabilidade de “testes autênticos”.

A *confiabilidade* de um item avaliativo pode ser definida como a medida em que os resultados de um teste são consistentes e acurados – o quanto o teste é eficiente na atribuição de valor a esses resultados. Se examinandos com nível de proficiência similar alcançam resultados semelhantes entre si num dado teste, pode-se dizer que sua confiabilidade é alta. Mas se esses resultados são muito discrepantes, ou se não é possível através do teste identificar os diferentes graus de proficiência existentes entre examinandos, dizemos que a confiabilidade do teste é questionável. Para uma discussão do termo, ver Shohamy (1985a, p. 70-73), onde a autora define e indica diferentes tipos de confiabilidade, relacionadas à consistência interna do teste, aos examinadores e aos processos de aplicação e correção dos exames. Nesse sentido é que para Lewkowicz (op. cit.), tantos e tão variados são os usos de língua que seria difícil delimitar o que pode ser tomado como parâmetro na definição do que é autêntico numa avaliação.

¹⁵ Retomo a discussão sobre proficiência na página 85. Para uma discussão mais detalhada, ver Scaramucci (2000).

Outro exemplo de pesquisas que vêm se detendo sobre a questão da integração de habilidades em exames de proficiência são os estudos dedicados ao novo formato do TOEFL (*Test of English as a Foreign Language*), desenvolvido pelo ETS (*Educational Testing Services*). O TOEFL, em sua versão antiga, era conhecido pela sua concepção mais estruturalista de língua e formato de múltipla escolha. As modificações introduzidas no “Next Generation TOEFL” visam a maior integração de habilidades e, poderíamos dizer também, maior proximidade ao extremo integrado da definição de integração proposta por Lewkowitz (1997) e descrita anteriormente, uma vez que essas mudanças refletem uma necessidade de maior autenticidade identificada pelos seus formuladores (cf. Trites & McGroarty 2005, p. 175).

Um dos estudos que se concentram sobre a questão é o de Lumley & Brown (2004), que analisa as perspectivas de examinandos e examinadores sobre o formato integrado das tarefas de escrita do exame. Um dos pontos que chama a atenção no estudo é o uso dos termos *independente* e *integrado* (*independent* e *integrated*) para a referência a tarefas que abarquem apenas uma macro-habilidade e tarefas que envolvam mais de uma macro-habilidade, respectivamente. O uso dessa terminologia é interessante para o presente estudo, uma vez que um certo desconforto parece se instalar quando usamos o termo *isolamento* de habilidade para referência a testes que exigem somente a leitura, talvez pelo fato de a tradução em português adotada na área de Lingüística Aplicada para *discrete point* ser *isolamento*. O sentido em que aqui se usa *isolamento*, no entanto, não é o mesmo que pode ser inferido do termo mais estreito *discrete point*, por esse se referir mais usualmente a um tipo de avaliação muito mais específica do que a simples separação de uma macro-habilidade – a separação de um elemento lingüístico de praticamente todo o contexto onde ele naturalmente se insere.

A questão terminológica, aliás, é um ponto importante a ser discutido. Visto que a maior parte da fundamentação teórica da pesquisa é composta por estudos produzidos em língua inglesa – não pela desconsideração de estudos realizados em contexto nacional, mas pelo fato de no Brasil praticamente *não* existirem estudos sobre a questão da integração de habilidades e autenticidade na avaliação –, a terminologia usada tem que ser importada do

inglês e, alguma vezes, não encontra correspondentes satisfatórios em português. A questão assinalada com relação a *isolamento/ discrete point/ independente* é uma delas.

Considero pertinente adotar o termo *independente* neste estudo, em detrimento de *isolado*, que já demonstrou produzir certa confusão de sentidos sobre exatamente que tipo de isolamento o trabalho se refere. *Independente*, aqui, equivale ao *independent* de Lumley & Brown (op. cit.) na referência à tomada de apenas uma macro-habilidade (leitura, escrita, compreensão ou produção oral) na avaliação, o item revelando-se integrado em termos contextuais ou não – isto é, tendendo para o extremo integrador que busca autenticidade da avaliação ou não; *independente* refere-se apenas às quatro macro-habilidades lingüísticas tomadas independentemente na avaliação, sem se levar em consideração o grau de autenticidade e relação com o uso real de língua que demonstre.

Com relação a outra questão terminológica aqui presente, o uso de *integrativo* e *integrado*, serão usados os termos *menos integrador* e *mais integrador*, pois, como fazem parte de um mesmo campo semântico, faz sentido dizer compartilham características em maior e menor grau. Além disso, esse uso é consoante com a idéia de que existem “níveis” de integração, também sugerida pela noção proposta por Lewkowicz (1997) de *integrativo* e *integrado* como extremidades de um continuum da integração na avaliação.

Tentei demonstrar, em toda a seção 2.1, a complexidade envolvida na integração e isolamento de habilidades, e a necessidade da adoção de um arcabouço teórico mais amplo, que dê conta, ou pelo menos tente de forma mais eficaz, de explicar aspectos relativos às visões de uso mais e menos real de língua na formulação e aplicação de avaliações em LE. Passo agora a discorrer sobre conceitos que acredito serem essenciais para o alcance dos objetivos de reflexão sobre a integração e isolamento de habilidades sob a perspectiva dos examinandos. Na seção 2.2. *Autenticidade na avaliação*, trato desse arcabouço teórico mais amplo a que me referi acima, que busca discutir de forma mais profunda o tópico da pesquisa, e na seção seguinte, 2.3. *Perspectiva dos examinandos*, enfoco mais particularmente a importância da investigação das percepções dos examinandos sobre a integração ou não de habilidades na avaliação e sobre as possíveis relações de autenticidade segundo a perspectiva deles.

2.2. Autenticidade na avaliação

A *autenticidade* é um dos traços que, segundo Bachman & Palmer (1996), constrói a utilidade de uma avaliação. Se o propósito de uma avaliação é fornecer uma amostra representativa do conhecimento ou do desempenho de um indivíduo relativo a um determinado construto, o mais apropriado é um método de avaliação que busque reproduzir as condições nas quais esse construto é habitualmente utilizado. Ou seja, o teste deve procurar reproduzir uma situação autêntica de maneira a avaliar mais precisamente a capacidade do indivíduo. A *situação autêntica*, no caso, configura o que eles chamam de *tarefa de uso da língua alvo* (*target language use (TLU) task*), que se refere ao propósito com que um domínio específico de uma língua (LM ou LE) é usado (cf. Bachman & Palmer *ibid.*, pp. 44-45). A Figura 2 sintetiza a concepção de autenticidade dos autores, definida como *o grau de correspondência entre características de uma tarefa avaliativa de língua e as características de uma tarefa de uso da língua alvo*.¹⁶

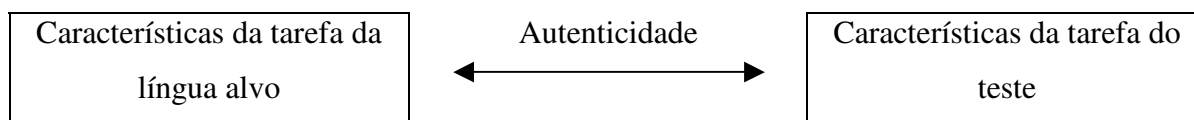


Figura 2: Autenticidade (Bachman & Palmer, 1996, p. 23)

Embora a definição de autenticidade seja clara, o modo como a autenticidade se constitui dentro dos mais variados testes não é absoluto. Visto que as características das tarefas de uso da língua alvo dependem de variáveis como o propósito de uso da língua, o contexto em que a tarefa é realizada, os participantes envolvidos, o nível de proficiência necessário, entre outros fatores particulares a cada situação, a autenticidade na verdade está muito sujeita às percepções que os envolvidos na situação de avaliação têm sobre o que é uso real de língua e os elementos que a compõem.

Esse é apenas um dos aspectos que torna a busca da autenticidade nas avaliações em LE complicada e no qual muita atenção e cuidado são necessários. Embora muito entusiasmo seja devotado à questão, a avaliação autêntica sendo algumas vezes

¹⁶ Citação original: “We define authenticity as the degree of correspondence of the characteristics of a given language test task to the features of a TLU task.” (Bachman & Palmer, 1996, p. 23.)

encarada como a panacéia de todos os males no campo da avaliação – e, conseqüentemente, no do ensino e aprendizagem de línguas –, muitos fatores devem ser levados em consideração na implantação de exames e testes que se pretendam “autênticos”. Além disso, como ressaltam Cumming & Maxwell (1999), interpretações equivocadas do que significa autenticidade podem levar a práticas de validade e confiabilidade duvidosas.

Em artigo que procura discutir os efeitos que diferentes interpretações do conceito de autenticidade nos meios educacionais produzem na construção de avaliações (não apenas na área da linguagem), os autores descrevem problemas encontrados em muitos contextos em que a busca por essa autenticidade acabou levando, na realidade, a uma redução da qualidade do produto dessas avaliações. Itens mal formulados (como, por exemplo, itens que tentam “camuflar” uma avaliação de conteúdo específico sob uma máscara de aplicação de conhecimento conceitual à vida real), escolhidos segundo critérios obscuros ou não apropriados para a situação de aplicação da avaliação, e a falta de compatibilidade entre as noções de autenticidade de examinandos e examinadores são alguns dos problemas elencados no artigo. Assim como Bachman & Palmer (op.cit.), os autores sugerem que a relatividade da autenticidade é característica a qual deve ser prestada primordial atenção para a garantia da validade da avaliação.

2.2.1. A autenticidade e os “testes diretos”

As considerações feitas até aqui sobre a autenticidade vão ao encontro do que Low (1985) propõe como ponto de partida para a validação de testes que aspirem à semelhança com a língua usada em situações reais. O autor usa o termo *testes diretos* (*direct tests*) para se referir ao que hoje mais comumente chamamos de testes autênticos. O termo “direto” diz respeito justamente ao modo de avaliação que esse tipo de instrumento propõe – debruçar-se de maneira direta sobre o construto a ser avaliado, exigindo a demonstração da capacidade de proficiência oral através de um teste que recolha amostras desse construto, e não através de um teste de gramática, por exemplo. O que hoje soa até certo ponto óbvio não o foi durante os primórdios da aplicação de avaliações de línguas, e não o é em alguns contextos de abordagem mais estruturalista, em que o conhecimento sobre a gramática ainda é usado como termômetro único de mensuração do saber uma

língua – um exemplo de teste indireto. *Testes diretos*, portanto, seriam testes em que situações de uso de língua são avaliadas *diretamente*, ou seja, as tarefas avaliativas são dispostas da maneira que habitualmente são feitas no uso cotidiano (por exemplo, solicita-se que o examinando escreva uma carta para se avaliar sua habilidade de escrita). Já testes indiretos seriam o oposto: o teste propõe uma situação que não corresponde diretamente ao modo como a língua é usada habitualmente (por exemplo, a avaliação da habilidade de escrita através de tarefas de múltipla escolha)¹⁷.

Low (ibid., p. 163) aponta como problema na validação de testes diretos a falta de balizas que definam o que é uma situação autêntica de uso de LE. Se pensarmos na variedade de situações em que a língua atua como instrumento de ação no mundo, podemos ter uma idéia do quão árduo seria chegar a tal definição para cada uma dessas situações, e, no mesmo sentido, do quão complexo seria estabelecer um critério justo de validação de tais testes. O problema, porém, persiste, e o ideal seria que os examinadores se preocupassem em definir, pelo menos num universo reduzido, os padrões de desempenho que se espera numa situação real e adaptá-los para a situação de avaliação.

Questões relacionadas a características de avaliações de LE, como a sua validade, são sempre pontos controversos e suscitam extensa discussão. É senso comum na área de avaliação no contexto de ensino de línguas, por exemplo, afirmar que um teste ganha em validade o que perde em confiabilidade¹⁸. Por exemplo, ganha-se em confiabilidade com testes mais objetivos¹⁹, pois há uma garantia maior de que os resultados seguirão padrões semelhantes quando aplicados em públicos parecidos e em circunstâncias similares; a validade desses testes, no entanto, é questionada, pois a avaliação pode não se conformar às circunstâncias em que o construto avaliado é habitualmente acessado. Por outro lado, avaliações de abordagem mais comunicativa possuiriam validade mais alta,

¹⁷ Uma discussão mais aprofundada sobre os testes diretos e indiretos pode ser encontrada em Clark, J.L.D. (1975), um dos primeiros a usar essa terminologia na área de avaliação de língua. Shohamy (1985a, p.5) também traz a definição de testes diretos e indiretos, e Bachman (1990, p.287) tece críticas sobre os mesmos.

¹⁸ Os conceitos de validade e confiabilidade são brevemente discutidos nas páginas 39 e 46, respectivamente.

¹⁹ O termo “objetivo” é aqui usado no sentido de testes que prevêm uma grade de correção mais fechada e menos sujeita a variações subjetivas – como os testes de múltipla escolha, por exemplo, típicos da fase psicométrica-estruturalista da avaliação (ver Shohamy 1985a, p. 4).

porém confiabilidade mais reduzida, dada a natureza diversificada dos produtos da avaliação, que torna mais difícil a padronização da atribuição de valor.

A princípio, um teste que se propõe autêntico é mais válido por sua própria natureza. O estudo de Low (ibid.), no entanto, já questiona essa premissa com argumentos significativos, como os que discuti acima, o que lança questões a serem consideradas quando tratando da validade de avaliações autênticas. Refletir a variedade dos usos de língua real nos testes, ao mesmo tempo em que tornaria as avaliações mais autênticas, é também um dos pontos sobre os quais trabalho metuculoso é necessário, pois pode se tornar uma desvantagem do instrumento avaliativo caso não seja zelosamente cuidado, pois mais variedade de língua implica mais variáveis a serem consideradas na correção. Essa mesma questão parece refletir parte das reservas em relação à confiabilidade: a variedade de uso da língua é tão grande que garantir menos subjetividade a essas avaliações não é um problema de fácil solução, como já afirmou Shohamy (1985b, p. 53). A autora ainda acrescenta que a alta confiabilidade de testes não-autênticos se dá justamente pela sua falta de validade, pela não preocupação com a diversidade da língua em uso, e identifica variáveis que atuam no distanciamento da “língua de uso real” da “língua desempenhada em situações de teste”, tais como os objetivos das interações, os participantes envolvidos, o cenário de desempenho, o assunto e o tempo estipulados. Ela ainda indica a consciência de que não existe a avaliação de LE “totalmente autêntica” como uma das primeiras premissas a que os examinadores devem atentar – pois a língua usada numa avaliação de LE sempre será linguagem de avaliação em algum nível, por mais que tente se aproximar da linguagem de dia-a-dia. Para Alderson (2000, p.53), a questão não é quão real a linguagem do teste é, uma vez que a leitura realizada numa situação de teste é um evento tão autêntico e presente no cotidiano quanto a leitura realizada em outros tipos de situações – a questão seria, pois, averiguar o quanto os resultados mostrados na situação de teste permitem que se possa generalizar sobre o desempenho lingüístico em outras situações de uso da LE.

No mesmo sentido, Widdowson (1998), um dos primeiros a tocar na questão da autenticidade no ensino de LE (v. também Widdowson 1978, 1979), estabelece a diferença entre “LE autêntica” e “LE apropriada” para o ensino. Segundo ele, muitas vezes os exemplos reais de língua não são satisfatórios para o ensino devido ao fato de toda e

qualquer amostra estar inexoravelmente intrincada ao contexto do qual faz parte e, por isso, perder parte – se não todo – seu significado nessa transferência de um ambiente real para um ambiente de sala de aula, ou sua inserção num livro didático. Para ele, é a comunidade lingüística que “autentica” a língua, atribuindo sentidos a determinados elementos dentro de um determinado contexto compartilhado. Longe de advogar, no entanto, que a busca da linguagem mais autêntica possível não é necessária no ensino de LE, ele sugere, ao contrário, que o professor busque maneiras de contextualizar a língua trazida para sala de aula, para que os próprios alunos atribuam sentidos a ele, “autenticando” esse insumo.

2.2.2. A relatividade da autenticidade

Fazendo um paralelo das considerações de Widdowson (1998) sobre o ensino de LE com sua avaliação, pode-se dizer que a linguagem usada numa avaliação também não tem condições de ser totalmente autêntica, pois o próprio ato de retirar um texto do seu contexto de produção já o tornaria “não-autêntico”. É possível dizer, então, que a autenticidade de uma avaliação de LE é relativa em dois sentidos: i) no sentido de que depende da percepção do público a quem se dirige sobre o que é uso autêntico de língua e ii) no sentido de que não constitui um tipo de avaliação “absoluto”, mas que seus itens podem ser mais e/ou menos autênticos, se aproximando ou se distanciando mais de um conceito virtual de autenticidade – daí ser mais apropriado falar em itens *mais* e *menos autênticos*, e não simplesmente *autênticos* ou *não autênticos*. A Figura 3 procura representar essa idéia, em que “Item A” simbolizaria um item mais autêntico e “Item B” um item menos autêntico em relação a esse conceito do que seria totalmente autêntico e do que seria totalmente não-autêntico.

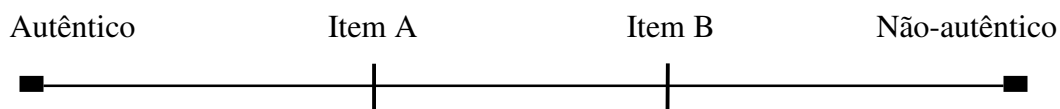


Figura 3: Um dos sentidos da relatividade da autenticidade

A definição de onde nesse continuum um item deve recair irá variar de acordo com fatores como para quem a avaliação é dirigida, quais seus objetivos e qual a visão de

língua que a subjaz. Tome-se o exame CELPE-Bras (*Certificado de Proficiência em Língua Portuguesa para Estrangeiros*), por exemplo. O CELPE-Bras²⁰ é um exame de proficiência em português para falantes de outras línguas certificado pelo MEC (*Ministério da Educação do Brasil*). Ele é exigido por diversas instituições educacionais e para validação de diplomas estrangeiros no Brasil. Está voltado para averiguar se os candidatos são capazes de realizar tarefas de uso do português com as quais eles se deparariam vivendo, trabalhando e estudando no Brasil. Devido a essas características, o CELPE-Bras se constitui como um exame que procura avaliar habilidades de forma integrada e assemelhar seus itens o máximo possível às tarefas de uso do português no cotidiano, usando textos orais e escritos autênticos (não construídos especialmente para o exame) e propondo tarefas integradas e com um objetivo verossímil, com um interlocutor definido, propósito comunicativo e gênero especificado. Pode-se dizer que os itens do CELPE-Bras procuram situar-se o mais próximo possível ao extremo “Autêntico” na faixa da Figura 2 (página 49).

O outro sentido de relatividade a que me referi, o das percepções sobre a autenticidade do uso de língua, tem sua importância ressaltada se pensarmos no que habitualmente é tido como mais autêntico em termos de integração de habilidades. Como já resaltei na subseção 2.1.2, a integração de macro-habilidades à primeira vista parece corresponder mais à idéia do modo como a língua é usada no cotidiano. Mas, analisando com mais cuidado esse quadro – e, o mais importante, tendo em mente que trato aqui de LE e não de LM –, as situações de uso de língua em que o isolamento é que constitui uso autêntico são comuns, como no caso de pessoas que usam a LE somente para ler, não tendo necessidade de escrever ou conversar usando a LE. Ou seja, a prática social da LE a que estão expostos orienta a percepção que eles possivelmente terão de um teste de habilidades integradas como pouco autêntico, ou pelo menos não tão autêntico como um teste isolado de leitura.

Lewkowicz (2000) preocupa-se com essa questão num estudo em que procura investigar as percepções dos examinandos sobre a autenticidade de dois testes, um de formato mais integrador e outro de formato independente. Seus resultados, de maneira

²⁰ Para uma discussão mais aprofundada dos princípios que norteiam o exame, ver Scaramucci (2004b) e Scaramucci & Rodrigues (2004).

geral, apontam para a idéia já esboçada acima da relatividade da autenticidade, e sugerem questões a serem respondidas em pesquisas futuras, sobre que fatores atuam na percepção da autenticidade e o aprofundamento da relação entre percepção e desempenho no teste. Ela ainda chama atenção para o fato de que muitas vezes as perspectivas dos examinandos podem diferir muito das de examinadores, o que poderia trazer conseqüências outras que não o sucesso na realização dos itens. Embora sugerido em muitos outros estudos sobre autenticidade, alguns dos quais citados neste trabalho, o estudo de Lewkowicz (ibid.) ressalta sobremaneira a importância que a investigação das perspectivas dos examinandos têm para a compreensão sobre o que entendemos por avaliações de língua mais autênticas e seu adequado desenvolvimento.

A identificação dessa ligação forte entre autenticidade e percepções dos examinandos, na realidade, retoma as hipóteses iniciais desta pesquisa. O fato de uma maior ou menor integração de habilidades estar relacionada a determinadas visões de língua e avaliação é algo que depende dos diversos elementos que compõem o contexto dessa avaliação, assim como a autenticidade dessas avaliações ser relativa à perspectiva que se toma para se olhar para a questão. Como já apontei anteriormente, a integração de habilidades está muito ligada à questão da autenticidade, e a autenticidade, por sua vez, está ligada às diferentes percepções dos envolvidos na avaliação, pois é a percepção de cada um deles que determina o que é ou não autêntico, não havendo uma entidade absoluta “autenticidade” aplicável a qualquer situação – ou seja, ela é *relativa* às diferentes percepções. As três questões estão muito amarradas teoricamente, os conceitos muito próximos e, por isso, estudar as perspectivas dos examinando passou a cada vez mais parecer muito significativo.

Tendo demonstrado os motivos que fazem da perspectiva dos examinandos questão importante para este estudo, discorro a seguir mais detalhadamente sobre esse tópico, e mais especificamente sobre essa perspectiva no que se relaciona à autenticidade das avaliações de LE.

2.3. A perspectiva do examinando

Muitos são os fatores que atuam na composição da percepção do examinando sobre o teste. Uma delas é a própria concepção que o examinando tem sobre o que é uma avaliação de LE. Isso irá, por exemplo, influenciar o quanto ele acredita que aquela é uma avaliação válida. O conceito de *validade de face* está envolvido nessa questão, e embora trate de um aspecto do conceito de validade muitas vezes negligenciado, por referir-se a uma espécie de validade mais “aparente”, atribuída pelo leigo, considero-o um fator importante a ser levado em consideração. A percepção que o examinando tem sobre o teste pode influenciar o modo como ele se prepara para realizá-lo, que conteúdos e/ou habilidades ele acredita que são importantes para realizá-lo bem, que aspectos priorizará e que estratégias desenvolverá, entre outros elementos. Esses aspectos, por sua vez, podem atuar no desempenho do examinando e nos resultados que ele apresentará ao final da avaliação.

Assim como a concepção sobre avaliação do examinando, a sua concepção sobre uso da língua e especificamente da leitura, sua motivação para realização do teste, seu nível de proficiência, entre talvez diversos outros fatores ainda por serem investigados, agem na composição da percepção de um instrumento avaliativo. Estudar e buscar compreender esse intrincado pode nos levar a pensar sobre questões que muitas vezes não são levadas em consideração quando tratamos de avaliações, como, por exemplo, se as expectativas dos diferentes envolvidos no processo de avaliação, os *protagonistas* do processo (*stakeholders*), como examinandos, corretores, elaboradores, agências e instituições que se utilizam dos testes etc., correspondem-se minimamente.²¹

Esse tipo de questão vai além de buscar compreender como o examinando percebe ou interpreta o enunciado de um item, embora isso faça parte da relação que se estabelece entre examinando e examinador; busca-se, em primeiro lugar, investigar em que medida as prioridades de examinandos, examinadores e outros envolvidos no processo da

²¹ O termo *stakeholder*, na literatura de avaliação de língua, refere-se a todas as pessoas e/ou instituições que fazem parte do processo de avaliação em algum nível: os responsáveis pela formulação e correção, os examinandos, e as instituições ou pessoas que usam os resultados das avaliações para algum fim seletivo, como uma empresa num exame de seleção para um emprego, uma instituição educacional para o ingresso num programa de estudos, uma agência financiadora para distribuição de bolsas de pesquisa, entre outros usos. Neste trabalho, adotaremos, como equivalente a *stakeholder*, o termo *protagonista*.

avaliação estão localizadas no mesmo nível e no mesmo campo. A falta dessa correlação acarreta problemas tais como examinandos tentando traduzir textos em LE o mais fielmente possível em avaliações de leitura que, para os examinadores, buscam avaliar habilidades outras, tais como a capacidade de síntese ou inferência, por exemplo. O problema talvez resida não na incapacidade do examinando em realizar sínteses ou inferências quando lendo em LE, mas na sua concepção e percepção de avaliação de leitura como espaço em que a tradução é evidência de compreensão do texto e de proficiência lingüística.

Na questão da autenticidade encontramos aspectos semelhantes. Como dito anteriormente, a autenticidade é uma característica que é tomada como importante e “positiva” numa avaliação bem feita e de princípios comunicativos. Devemos, no entanto, estar atentos à adoção desse pressuposto sem os devidos cuidados e esclarecimentos sobre seu construto. Não apenas no campo de avaliação de LE, mas na avaliação em geral, existe a tendência em atribuir valor à autenticidade em si, sem a atenção aos outros elementos constituintes da avaliação, e algumas vezes a tal busca pela “língua de uso real” não necessariamente é o melhor caminho na construção de testes válidos e confiáveis. Como afirma Widdowson (1998), nem sempre o uso de língua que se faz nos ambientes “autênticos” é também o tipo de linguagem mais apropriado a ser empregado no ensino de uma LE. Assim, questionar em que condições, e para quem, a autenticidade constitui uma característica desejável e proveitosa numa situação de avaliação é relevante e merece mais atenção da parte dos envolvidos nos processos avaliativos existentes em contextos escolares e externos.

A autenticidade, constituindo um conceito relativo principalmente ao contexto em que é tomada, ao público a quem é dirigida a avaliação, e às visões de língua e avaliação que esse público traz, faz com que imperativo seja se pensar sobre diferentes percepções desse mesmo público. No caso de professores de línguas e responsáveis pela formulação e correção de avaliações de línguas em larga escala, por exemplo, as percepções sobre um teste estarão calcadas em fatores como experiência anterior no lidar com avaliações de língua, conhecimento técnico da área e concepções e pressupostos teóricos construídos por meio de estudos, pesquisa e empirismo. Esses elementos é que orientarão a perspectiva que adotam frente a uma avaliação, e a análise que realizam sobre

a validade de um item carrega todos esses elementos. Essas percepções orientam ainda a formulação, aplicação e correção de avaliações de línguas que esses profissionais possam empreender.

Já no caso das pessoas que se submetem a avaliações de línguas, encontram-se indivíduos que baseiam suas percepções sobre os testes também no que acreditam ser adequado e válido, mas obviamente a perspectiva adotada difere – ligeiramente ou não – da perspectiva dos especialistas da área. Dessa forma, o julgamento que fazem sobre a possível validade de um item é de caráter mais superficial, baseada na aparência do item, e não necessariamente em traços que o professor de língua ou estudioso do campo da avaliação consideraria. A validade atribuída pelo leigo é o que na literatura da área de avaliação de línguas é denominada *validade de face* – um tipo de validade “aparente”, que diz se o item *parece* ser válido, ou seja, se parece ser um instrumento adequado ao que propõe avaliar, baseado apenas em seu aspecto mais superficial, sem a análise técnica empreendida dentro dos rigores necessários para a validação feita por um especialista da área.

2.3.1. A validade de face

Como já afirmei na página 16, ao introduzir o conceito de *validade de face*, ela merece mais atenção por parte dos estudiosos como característica de um exame. Pode-se dizer que a validade de face funciona como uma medida da aceitação de um exame e, se investigada adequadamente, pode demonstrar se os objetivos dos formuladores estão sendo atingidos junto aos examinandos e, também, o que os examinandos acreditam ser importante na realização do teste.

Nevo (1985) critica a negligência e a trivialidade com o que o conceito é tratado na área de avaliação e defende que estudos mais engajados sobre a validade de face dos exames devem ser feitos. O autor sugere que dois testes equivalentes em termos de construto avaliado podem ser empregados de formas diferentes devido à diferente validade de face de cada um – o teste com maior validade de face seria mais eficiente como avaliação devido à sua “aparência” mais válida convidar a um envolvimento maior dos examinandos em sua realização. O autor descreve métodos de mensuração da validade de

face de uma avaliação e aponta como idéia valiosa a divulgação de dados sobre a validade de face de exames em seus próprios manuais, como evidência séria da qualidade do exame. Como argumentos em favor desse gesto, ele enumera alguns fatores que podem promover o aumento de um tipo de atitude “positiva” em relação ao exame: maior envolvimento e motivação dos examinandos, a atração de potenciais candidatos, a redução da insatisfação com relação a desempenhos fracos, o convencimento de protagonistas tais como empregadores, administradores e responsáveis políticos na implementação de tais exames e a melhora na questão das “relações públicas” do exame, uma vez que a validade de face seria a ponte entre os responsáveis pelo exame e o público em geral, incluindo setores de mídia e jurídicos, por exemplo (cf. Nevo, op. cit., p. 288)

Zeidner & Bensoussan (1988, p.113), no mesmo sentido, ressaltam que investigar as percepções dos examinandos sobre as avaliações pode ajudar examinadores a identificar pontos problemáticos na administração de exames, que vão desde os itens em si até o limite de tempo para realização do exame. As autoras ressaltam ainda que dar a oportunidade aos examinandos de expressar suas opiniões sobre a avaliação feita pode ser um caminho para reduzir a ansiedade e apreensões com relação ao exame e, assim, tornar a experiência mais rica e significativa.

Muitas críticas sobre a validade de face versam sobre aceitar seu aspecto “atrativo” como evidência suficiente do valor e qualidade da avaliação, sem a devida verificação se o teste realmente dá conta de avaliar os construtos e conteúdos a que se propõe. Para esses teóricos, a validade de face se refere a um tipo de credibilidade apenas superficial, e estaria mais ligada a aspectos de aceitação social do teste do que a características técnicas que permitissem dizer se a avaliação é adequada ou não. (v. Cattell 1964, Ingram 1977).

Embora o argumento acima seja pertinente, é arriscado menosprezar a importância da validade de face baseada na premissa de que sua superficialidade constitua uma falha que a desacredite. Como já afirmei anteriormente, a aceitação social de um teste é um elemento significativo no complexo que envolve o processo de avaliação. Das inferências que podem ser feitas sobre a capacidade lingüística de um sujeito a partir de um teste são tomadas decisões que podem afetar indivíduos, grupos e instituições em diferentes

graus. Muitas dessas decisões não dependem de especialistas da área de avaliação de línguas, mas podem caber a outros envolvidos que participem do processo, tais como responsáveis pela seleção e ingresso em programas educacionais ou profissionais, pela atribuição de bolsas de estudos, pela formulação e manutenção de currículos entre outros. Se o teste “parece” válido para esses protagonistas, ou seja, se sua validade de face é tal que permita que ele seja aceito por esses não-especialistas, pode-se dizer que este tipo de validade também constitui uma característica importante do teste.

Obviamente, deve-se concordar com os críticos da validade de face no que se refere a tomá-la como evidência única da qualidade do teste. A “aparência” do item é apenas um dos aspectos de sua validade, e é necessário estar atento a outras características no processo de validação de um teste, para se garantir que além de parecer avaliar adequadamente um construto ou conteúdo, ele realmente busca fazê-lo, e/ou efetivamente o faz. A aceitação dos chamados “testes diretos” como válidos *a priori* é criticada por Bachman (1990, p. 287), que, embora reconheça a importância desse caráter mais aparente dos itens, afirma que questões cruciais sobre o conceito continuam sem serem respondidas, como as apontadas por Davies (*apud* Bachman *ibid.*). Este autor aponta que muitas vezes é difícil se chegar a um acordo sobre se um teste parece ou não válido e aceitável, pois a variedade de “leigos” é tão grande que abre um impasse: a opinião de qual “tipo de leigo” deve ser considerada? Mesmo a dos que apresentem uma perspectiva equivocada ou antiquada sobre aprendizagem e avaliação de línguas?

Provavelmente não existe solução simples para essa questão, talvez por não ser possível tratar da validade de face do mesmo modo que se trata dos outros tipos de validade. Sendo um conceito essencialmente sujeito às percepções que se tenha sobre a avaliação, não é possível estabelecer um método de construção de testes que garanta sua validade de face independente de variáveis como o público a quem se dirige o teste, o tipo de domínio da LE que se busca abarcar, os objetivos da avaliação, entre outros pormenores.

Além disso, embora seja plenamente válido questionar, como os autores acima, a possibilidade de se saber quem achará que teste é aceitável e em que condições, devido à grande variedade de contextos de aplicação de exames de línguas, isso não pode ser motivo para que teóricos da área da avaliação de línguas negligenciem o problema que ele

obviamente representa. Ao contrário, buscar cada vez mais elucidar a questão de modo a partir para práticas avaliativas mais conscientes de seu alcance e, principalmente, de suas limitações, é imperativo. E, ligado a isso, um ponto que carece de maior atenção é justamente a heterogeneidade de percepções sobre as avaliações de língua segundo as perspectivas dos diferentes protagonistas. Como já apontei anteriormente, investigar o quanto se correspondem as expectativas de examinandos e examinadores sobre os resultados dos exames e as inferências sobre a capacidade lingüística dos examinandos que podem ser feitas baseadas nesses resultados é primordial para a satisfação de ambas as partes. Dessa forma, examinadores estariam mais seguros de que seus testes alcançaram seu objetivo – seleção, colocação, diagnose etc. – e examinandos de que o processo avaliativo foi justo e correto.

2.3.2. Diferentes contextos da autenticidade na avaliação

Não apenas com relação a exames externos de LE existe a importância da investigação dessas percepções, mas também em contexto de sala de aula. Zeidner (1988, p. 231) afirma que a probabilidade de que exista cooperação dos examinandos, bom relacionamento entre professor e aluno e melhora na motivação para fazer o teste se os examinandos o perceberem como instrumento válido e justo, minimizando reações adversas. A autora sugere ainda que investigar essas percepções pode ser útil no repensar práticas avaliativas e mudanças e remédios de falhas nelas presentes.

O estudo de Moni et. al. (2002) corrobora as considerações feitas por Zeidner (op. cit.). As autoras investigaram, num estudo longitudinal, as percepções de alunos (iniciando no que seria o equivalente ao nosso ensino médio) sobre a avaliação de LM empreendida em sala de aula e constataram que elas variavam muito – talvez devido às diferentes experiências de avaliação que cada um deles já trazia de sua vida escolar – e que, enquanto em alguns casos durante o período de estudo elas permaneciam as mesmas, em outros elas mudaram. Isso poderia ser explicado talvez pelo tipo de experiência anterior com avaliações escolares que esses alunos traziam, se elas criavam expectativas que correspondiam ou não ao tipo de avaliação empregado no ano escolar em que o estudo foi realizado.

Outra constatação importante, feita a partir dos resultados dessa pesquisa, foi sobre as atitudes dos alunos com relação à avaliação, fortemente baseadas em suas percepções sobre ela. Quando tinham percepções mais positivas – quando atribuíam maior validade de face às avaliações, por exemplo, embora as autoras não usem explicitamente esse termo –, a tendência era que se envolvessem mais com a avaliação e com a construção de conhecimento que fazia parte dela, apresentando uma participação mais ativa em sala de aula. Isso destaca a importância da atenção a ser prestada às opiniões dos alunos sobre os instrumentos avaliativos e à formulação de métodos avaliativos que, se não totalmente, pelo menos tentem aproximar-se das expectativas dos examinandos sobre o que é um teste válido para garantir o envolvimento interessado deles na realização dessas avaliações. Outra alternativa é o trabalho do professor na demonstração aos alunos da validade do instrumento proposto, modificando uma percepção equivocada que eles possam ter. Em qualquer um dos casos, a importância da atenção às perspectivas dos examinandos mostra-se presente.

Pode-se dizer que o tipo de atitude devotado à avaliação, determinado parcialmente pelas percepções dos examinandos, faria parte do *efeito retroativo*²² (*washback* ou *backwash effect*) que essa avaliação exerce sobre eles. Efeito retroativo da avaliação se refere ao impacto potencial que ela exerce sobre os protagonistas do processo avaliativo e sobre processos e produtos de ensino e aprendizagem. Um dos estudos que relaciona percepções ao efeito retroativo de exames de LE é o de Bartholomeu (2002), no qual a autora busca investigar percepções, atitudes e motivações de alunos da terceira série do Ensino Médio sobre os exames de inglês de três grandes vestibulares do estado de São Paulo (USP, Unicamp e Unesp) e como elas influenciam a preparação para esses exames. Seus resultados sugerem que as visões e ações desses alunos estão muito ligadas ao contexto escolar que frequentam, e identificaram, assim como outras pesquisas já sugeriram, um efeito retroativo não determinista dos exames, mas dependente de outras variáveis, tais como as características individuais de cada aluno. Outro estudo que relaciona percepções sobre a avaliação ao seu efeito retroativo é o de Scaramucci (2004c) sobre o

²² Para uma discussão mais ampla sobre efeito retroativo, assim como uma exposição do estado da arte do conceito, ver Scaramucci (2004a).

exame CELPE-Bras, que investiga como as percepções de professores de português para estrangeiros e de candidatos ao exame influenciam suas práticas pedagógicas e a preparação que empreendem para o exame, respectivamente. Com relação às perspectivas dos candidatos (examinandos), recorte que interessa à minha pesquisa, a autora indica que as posturas da grande maioria deles é positiva, e suas respostas indicam que atribuem validade de face ao exame – o que favorece o efeito que o exame tem no modo como se dedicam à aprendizagem da LE.

Com relação a perspectivas sobre exames de proficiência de LE, outro estudo especialmente importante para o desenho da presente pesquisa é o de Lumley & Brown (2004) sobre a integração das habilidades de escrita e leitura no novo exame TOEFL (*Test of English as a Foreign Language*), desenvolvido pelo ETS (*Educational Testing Services*). O TOEFL, em sua versão antiga, era conhecido pela sua concepção mais estruturalista de língua e formato de múltipla escolha. As modificações introduzidas no “Next Generation TOEFL” visam a mais integração de macro-habilidades e o domínio no uso de micro-habilidades como realização de síntese e resumo. Como já mencionei antes, pesquisar as percepções de examinandos e examinadores pretendia investigar como eles interpretavam essas novas exigências introduzidas.

Uma das principais constatações desse estudo é semelhante ao que mencionei anteriormente sobre os motivos da mudança no formato do IELTS: muitos candidatos pareciam confusos sobre se os itens enfocavam leitura ou escrita, não demonstrando um entendimento claro da natureza integrada dessa parte do exame. As análises parecem sugerir ainda que o nível de proficiência desempenha papel importante nesse processo, pois candidatos com níveis mais baixos encontram mais dificuldade na compreensão dos textos e na seleção de informações relevantes, o que prejudica a qualidade do produto escrito que a tarefa exige.

Outro ponto importante do estudo é a relativa discrepância entre as perspectivas dos examinandos e examinadores: não há concordância em relação ao uso do insumo fornecido no componente de leitura para a realização da tarefa, pois os examinandos não parecem conscientes da quantidade excessiva de cópia que realizaram. Os autores apontam como uma das implicações do estudo a necessidade de que os examinandos tenham

preparação mais adequada para o exame e, mais importante, realizem uma reflexão crítica sobre o processo de avaliação.

A falta de correspondência entre as perspectivas dos diferentes protagonistas já foi mencionada aqui anteriormente, e mais uma vez se mostra como ponto importante na investigação das percepções dos examinandos sobre a autenticidade das avaliações. Como dissemos anteriormente, a autenticidade é uma característica tomada como importante e “positiva” numa avaliação bem feita e de princípios comunicativos. Devemos, no entanto, estar atentos à adoção desse pressuposto sem os devidos cuidados e esclarecimentos sobre seu construto, o que realmente está envolvido. Não apenas na avaliação de LE, mas em contextos de avaliação em geral, existe a tendência em atribuir valor à autenticidade em si, sem a atenção aos outros elementos constituintes da avaliação. Essa tendência é bem descrita no já citado trabalho de Cumming & Maxwell (1999, p.177):

“O conceito de avaliação autêntica foi adotado com entusiasmo pelos responsáveis pela criação e desenvolvimento de políticas e currículos educacionais e outros profissionais da área, e embutido na literatura de ensino e avaliação como uma característica desejável da avaliação. Claro, é tão difícil ser contra avaliação autêntica como o é ser contra torta de maçã e maternidade. É obviamente uma ‘coisa boa’. A alternativa, presumivelmente, seriam avaliações não-autênticas, e ninguém quer isso. As avaliações autênticas são claramente o caminho. Esse entusiasmo que abrange muitos setores educacionais levou a diferentes interpretações do que são avaliações autênticas. Embora seja natural que termos evoluam para o uso comum e adquiram sentidos diferentes, isso pode levar a teorias e práticas confusas. É tempo de considerar os diferentes sentidos atrelados a avaliações autênticas e como eles se relacionam a teorias de aprendizagem e sua avaliação. Esclarecimentos dessas questões devem levar a uma melhor prática.”²³

²³ Citação original: “The concept of authentic assessment has been embraced enthusiastically by policy-makers, curriculum developers and practitioners alike, and enshrined in the literature on curriculum and assessment as a desirable characteristic of education. Of course, it is difficult to be against authentic assessment as to be against apple pie and motherhood. It is obviously a ‘good thing’. The alternative, presumably, would be inauthentic assessment and nobody would want that. Authentic assessment is clearly ‘the way to go’. This enthusiasm across many educational sectors has led to many different interpretations of authentic assessment. While it is natural for terms to evolve into common use and acquire different meanings, this can often lead to confused theory and practice. It is timely to consider the different meanings attached to

Essa visão de avaliação autêntica como “obviamente uma coisa boa” pode estimular a busca pela formulação de testes cada vez mais próximos das situações reais de uso da língua. Muitos pontos controversos, no entanto, entremeiam esse processo, e no que se refere à perspectiva do examinando, podemos nos perguntar: será a autenticidade de um teste “obviamente uma coisa boa” também para eles? Ou mesmo: os examinandos têm um conceito do que é autenticidade na avaliação? Como eles percebem essa característica?

Como já citei anteriormente, quando introduzindo a questão da importância das percepções dos examinando, Lewkowitz (2000) realizou um estudo que buscava investigar que importância os examinandos atribuíam à autenticidade dos testes e como essa percepção afetava o desempenho deles. A autora conclui que a autenticidade não é percebida pelos examinandos necessariamente como algo importante, que permitiria avaliar suas capacidades lingüísticas de uma maneira melhor. Esses resultados são consoantes com o questionamento feito por Cumming e Maxwell (op. cit.) sobre se a avaliação autêntica é “obviamente uma coisa boa”.

Entretanto, o conceito de autenticidade sobre qual a autora firma suas conclusões não é “explícito”: os sujeitos expressaram suas percepções sobre dois tipos de testes, um tomado pela pesquisadora como mais autêntico e outro como menos autêntico, segundo os critérios que geralmente definem o nível de autenticidade do teste para teóricos da avaliação e examinadores; a partir dos dados recolhidos referentes a cada um dos testes, ela elabora uma reflexão sobre a percepção dos examinandos, porém seu questionário não incluía perguntas sobre a importância da autenticidade, e sim sobre os pontos fortes e fracos de cada um dos testes. A preferência, em muitos casos, pelo teste “não-autêntico” lança a seguinte pergunta: para quem os testes usados na pesquisa se configuravam como autênticos ou não autênticos? Também para os examinandos ou somente para os examinadores? Como já demonstrei anteriormente, a autenticidade é relativa, e o estudo parece tomá-la como conceito absoluto. Não teria justamente essa relatividade feito com que muitos não considerassem importante a pretensa autenticidade do teste – o fato de

simplesmente ela não constituir autenticidade para eles no mesmo sentido do que para aqueles que formularam os testes?

Busquei nas três seções anteriores desenvolver três tópicos que estão envolvidos no problema de pesquisa no presente estudo: *i)* o conceito integração de habilidades na avaliação, *ii)* autenticidade na avaliação e *iii)* perspectiva dos examinandos. Procurei mostrar como estão relacionados entre si, quais são as contribuições e lacunas nos estudos empreendidos anteriormente e que caminhos elas apontam como trajetória de pesquisa futura. Esses questionamentos orientaram o desenho metodológico do estudo, descrito no primeiro capítulo. A seguir, discorro sobre a aplicação dessa metodologia, tendo como fundamentação os conceitos discutidos neste capítulo, no cumprimento dos objetivos estabelecidos na pesquisa, descrevendo a coleta e apresentando a análise desses dados.

CAPÍTULO III

ANÁLISE E DISCUSSÃO DOS DADOS

Apresento, neste capítulo, a procedimentos de coleta e análise dos dados da pesquisa. Na sub-seção 3.1. *Estudo piloto* descrevo o estudo piloto – o perfil do sujeito, as etapas de aplicação, assim como os resultados obtidos. Retomo os alinhavos teóricos tecidos sobre o problema de pesquisa e as hipóteses sobre seu desenrolar na tentativa de construir vias de entrada na análise dos dados, assim como aponto elementos interessantes surgidos no estudo piloto. Na seção 3.2. *Coleta de dados*, discorro sobre as implicações que o estudo piloto trouxe para o desenho da pesquisa e os ajustes feitos nas abordagens teórica e metodológica, assim como os procedimentos da subsequente coleta de dados. Finalizo, a seção em 3.2.1. *Perfis dos sujeitos*, descrevendo os perfis dos sujeitos que participaram da coleta de dados.

3.1. Estudo piloto

O estudo piloto foi idealizado com intuito de refinamento do problema de pesquisa, buscando evidenciar possibilidades mais interessantes para sua discussão e possíveis falhas e limitações em sua estrutura metodológica. Essa primeira coleta, como já mencionei anteriormente, foi planejada na forma de uma entrevista semi-estruturada, tendo sido gravada em áudio e transcrita posteriormente.

O sujeito escolhido para essa primeira coleta de dados era do sexo feminino, com 28 anos de idade, e relatou a necessidade de ler textos em inglês-LE durante seu dia-a-dia em seus estudos de pós-graduação. Ela é aluna de mestrado na área de Biológicas numa instituição pública e vive/viveu a realidade de submeter-se a avaliações de leitura em inglês-LE no programa de pós-graduação que frequenta, que exige testes de proficiência no ingresso e no decorrer do curso. A examinanda ainda relatou ter realizado estudos de inglês em escolas de idiomas por 2 anos, em cursos que enfocavam o desenvolvimento não apenas

da leitura, mas também das outras três macro-habilidades relacionadas: escrita, produção e compreensão oral.

Na primeira parte da coleta de dados, a examinanda forneceu os dados acima e fez considerações sobre avaliações de leitura em LE a que já havia se submetido, sobre como geralmente realiza avaliações desse tipo, sobre sua percepção geral sobre eles e sobre diferentes tipos de itens sugeridos. Essa etapa durou cerca de 20 minutos.

Na segunda etapa²⁴, que durou cerca de 25 minutos, foi utilizada não somente a entrevista semi-estruturada, mas também um protocolo de retrospectiva²⁵. Nessa etapa da pesquisa, o método de retrospectiva foi utilizado para se averiguar se os dados assim coletados forneceriam maior riqueza interpretativa do que a entrevista semi-estruturada baseada na análise dos itens.

Alguns elementos que apareceram nessa primeira coleta foram²⁶:

a) a ansiedade gerada por avaliações:

EP009	Eu acho que sempre tem algum grau de ansiedade ... mas principalmente quando é de alguma coisa que eu não domino muito, como inglês, por exemplo.
EP013	Então, no caso desses testes, eu acho que foi importante ter contato com a prova antes, saber que tipo de prova era... eu acho que isso diminui a ansiedade , saber que tipo de prova que é... se são só perguntas em inglês, respostas em inglês... se a resposta é em português... eu acho que saber como é o tipo de prova é importante. E... que mais? Ah, e acho que, pra mim, acho que me dá um certo conforto estudar até a véspera, depois ter de preferência uma boa noite de sono, e poder fazer tranqüila o teste no outro dia.

b) o tempo estipulado como fator importante no desempenho nas avaliações:

EP007	Bom... eu lembro que a prova de entrada no mestrado, na verdade, foram quatro textos com perguntas todas em inglês e de múltipla escolha... e assim, pra mim, foi bastante exaustiva a prova, porque eu tinha 4 textos longos... assim de uma página toda... pra resolver em 3 horas. Então, pra mim, que não tenho muito tempo de estudo em inglês, eu acho que foi uma prova
-------	--

²⁴ Cf. APÊNDICE III.

²⁵ A *retrospecção*, técnica de relato verbal na qual o sujeito relata sobre o processo de leitura ou realização de uma atividade, é realizada após a realização do evento em investigação – após a leitura de um texto ou a realização de um teste, por exemplo. Difere da *introspecção*, que é o relato ocorrido durante o processo em questão, em que o sujeito *pensa alto* sobre o que está fazendo. Para uma discussão mais detalhada do método, assim como exemplos de usos em outras pesquisas, ver Cohen (1989).

²⁶ As partes destacadas referem-se a elementos que reforçam a identificação desses pontos na fala da examinanda; o uso de maiúsculas indica a ênfase feita pela própria examinanda, e o uso de reticências indica pausa ou hesitação; os códigos EPXXX se referem à localização do turno do sujeito do Estudo Piloto citado na extensão da entrevista; o código P refere-se aos turnos da pesquisadora e entre parênteses estão seus apontamentos.

	bastante cansativa. O tempo pra mim foi escasso. Eu consegui terminar a prova, mas eu acho que teria ido melhor se eu tivesse um pouco mais de tempo pra fazer a prova.
--	---

c) a dificuldade/facilidade percebida nas avaliações:

EP007	Agora, a segunda prova que eu fiz de inglês, que foi agora no meio do mestrado, foi mais tranqüila, eram 2 textos também, mas sobre o mesmo assunto... mais fácil que os anteriores que eu já tinha feito a prova, essa prova de entrada que eu falei... e achei que o texto era mais fácil . E além disso achei que o tipo de pergunta também era diferente, não era de múltipla escolha, eram questões abertas, a questão em inglês pro candidato responder em português. Então acho que esse tipo de prova é mais fácil do que a anterior que eu fiz também, que era de múltipla escolha, com tudo em inglês, perguntas e respostas em inglês. Essa que era com perguntas abertas, com perguntas em inglês e respostas em português eu acho que foi mais fácil .
EP027	Eu acho que é mais FÁCIL ... não sei se... pra MIM seria mais fácil ... ter um texto com questões abertas em inglês e poder responder em português.
EP032	Escrever em inglês, eu acho que... é... dificulta bastante , porque na verdade pra escrever em inglês... é, no meu caso, eu demoro bastante, porque... é... bom, se eu for querer escrever rapidamente, eu vou escrever certamente muitas coisas erradas, então teria que voltar pro texto e tentar corrigir meus próprios erros. Acho que isso demanda mais tempo, e acaba dificultando o teste .

d) o efeito retroativo do exame na preparação realizada de antemão:

EP012	Então, no caso da primeira prova que eu fiz, que foi oferecida pela Cultura Inglesa, eu peguei uma prova que tinha sido dada pela escola no ano anterior ... para um outro curso, mas eu sabia que minha prova seria feita naqueles moldes . Então eu estudei por ela, e vi como eu me saía naquela prova. Então eu tentei também da mesma forma fazer a prova , com o tempo que eu tinha também, sabendo que eram 3 horas, pra realizar a prova em 3 horas... e depois conferi os resultados pra ver como eu tinha ido. E pra segunda prova, eu também tinha várias... é, no caso essa segunda prova foi realizada pela Unicamp... aí eu tinha várias provas de anos anteriores , e eu também fiz a leitura dos textos... e das perguntas, pra ter mais ou menos uma noção mesmo do tipo de prova que era oferecido... e pra ver se eu conseguia realizar minimamente no tempo, né, que era dado pra realização da prova. Foi assim.
-------	--

e) capacidades, estratégias e conhecimentos lingüísticos necessários para a boa realização de avaliações de leitura em LE:

EP015	Eu acho que a questão do vocabulário e da gramática é importante , porque senão também você vai ter vocabulário mas vai escrever tudo errado, quando for escrever um texto em inglês... e pode ser que você também tentando, mesmo ler um texto, pode ser que você tenha um entendimento equivocado porque você não sabia um tempo verbal , ou alguma coisa do tipo.
EP017	(...) eu acho que o que ajuda, é, por exemplo, no meu caso pra fazer essas provas... entrar em contato com textos da área facilitou o conhecimento de termos que são técnicos e que são sempre empregados em todos os textos... então acho que facilita... acho que dá pra falar que facilita nesse sentido. Tem termos, expressões, palavras que se repetem ... termos técnicos mesmo, nesses textos.

EP018	“Bom, pra fazer a prova, como o tempo era muito escasso, tinha que entender o sentido global mesmo do texto, dar uma olhada nele... mas eu também procurava ler as questões... em alguns casos às vezes eu não entendia, por exemplo, um parágrafo do texto eu tinha uma questão... lendo a questão, entendendo a questão, acho que ela te direciona um pouco pro texto também. Então eu usava isso como estratégia também. Mas... é... acho que com o nível de inglês que eu tenho.. e com pouco tempo também... é impossível de ler o texto de cabo a rabo. Não dá, né. Então tem que ler mesmo... globalmente, se eu não entendi tudo não tem problema, dou uma olhada geral, e... e aí acho que é isso, e aí se tem alguma questão, no caso dessa prova, que eram questões-testes... se tem alguma questão que eu sei que tá naquele parágrafo que eu não entendi direito, aí eu vou procurar ficar um pouco mais de tempo em cima dele até entender o parágrafo de tal forma que não prejudique a resolução do teste.
EP019	Acho que eu vou ler o título... talvez eu leia as questões também primeiro, dos testes... agora, se não tiver teste junto, é... é acho que é isso, acho que eu vou ler o título e vou começar a ler o texto do começo.
EP024	Então, eu acho que sublinhando e no caso dessas duas provas que eu fiz, foi com uso de dicionário , quando eu procurava palavras que eu não sabia, eu também escrevia do lado.

f) percepção das diferenças entre a leitura realizada em avaliações e as leituras realizadas cotidianamente – ou seja, presença de estratégias para realização da prova:

EP018	Bom, pra fazer a prova , como o tempo era muito escasso, tinha que entender o sentido global mesmo do texto, dar uma olhada nele...
EP023	Acho que dependendo... é que depende , assim... se a intenção é responder o teste para aquela prova que você tá fazendo, eu procuro entender mesmo a pergunta , seja ela teste ou pergunta aberta... e buscar no texto. Já olhar pro texto meio que procurando aquilo lá.

g) percepções sobre a validade das avaliações:

EP032	Bom, esse teste de responder em português, eu fiz e achei que foi legal. Achei que foi... acho que... é um teste que avalia, se você conseguiu ler um texto e entender... e fica muito mais fácil pro candidato, que tá fazendo no idioma que ele domina.
EP033	(...) porque é um teste de leitura, né? Pra ver se você sabe ler, se você consegue ler um texto em inglês e entender. Então se é pra isso, se é pra ver se você consegue ler, entender aquele texto, eu acho que... é legal que seja um teste pra você responder em português. Se você responder corretamente em português significa que você leu o texto e entendeu.

h) percepções sobre o nível de autenticidade das avaliações:

EP034	(...) esse teste que eu também disse que acho que é mais interessante, que é ler textos em inglês e poder responder em português... porque eu leio textos em inglês... que são referências importantes pro meu trabalho... e eu vou fazer a referência dos mesmos em português... então eu tenho que ter entendido o que tá escrito lá, e aí vou escrever em português, né...
EP036	Porque esses testes que... é, por exemplo... da prova que eu fiz na Cultura Inglesa... esses testes em inglês... na verdade não se parece com atividade que eu tenho rotineiramente.

EP044	É, então a que eu mais considero parecida é isso, pegar um texto em inglês e poder responder em português. Afinal de contas é isso que eu vou fazer. Se eu tenho referência em inglês, que eu vou usar no meu trabalho, eu preciso entender e eu vou acabar escrevendo sobre esse texto que eu li em português, então eu acho que isso é bem mais parecido. E por consequência é isso mesmo que você falou, um teste que seria pouco parecido com atividades diárias , assim, da atividade profissional ou mesmo como aluna da pós-graduação... seriam esses testes de múltipla escolha.
-------	---

Os elementos que acabaram de ser apresentados dizem respeito a tópicos que surgiram na fala da examinanda como um todo, tendo sido retiradas de toda a extensão da entrevista. Algumas considerações mais pontuais sobre as tarefas realizadas estão transcritas a seguir, que vão no mesmo sentido das previamente citadas. Na segunda etapa da entrevista, a tentativa era ir mais a fundo nas questões relacionadas a autenticidade de cada uma das tarefas (Tarefa 1: múltipla escolha; Tarefa 2: resumo em LM; Tarefa 3: e-mail em LE):

EP052	(...) eu acho que a terceira tarefa do e-mail pode ser uma atividade bastante corriqueira na vida de muitas pessoas. Na minha não é. Não tenho ninguém com quem eu me comunique em inglês por e-mail. Mas eu acho que pode ser uma atividade corriqueira, mas não é o meu caso. Então considerando o meu dia-a-dia, eu acho que a segunda atividade (Tarefa 2) também é uma atividade que eu realizo mais no meu dia-a-dia, no meu trabalho.
EP053	P: Você acha que a 1, por exemplo, é uma atividade totalmente fora do seu... da sua vida real... ou parcialmente... Não, totalmente fora. Não faço nada desse tipo.
EP057	Se eu tivesse que classificar, dentre as tarefas, assim... a que eu escolheria pra eu fazer, que se parecesse mais com meu cotidiano , seria a 2, como eu já falei, depois o e-mail e por último a tarefa número 1, que é a de perguntas-teste.

Com relação à questão da validade atribuída pela examinanda, com relação à Tarefa 1 (questões de múltipla escolha), ela aponta pontos específicos que a fazem questionar sua validade (EP089:(...)*por exemplo, procurar palavras sinônimas.. não sei... não sei se sempre é uma boa... não sei se SEMPRE isso vai medir que você entendeu o que tá escrito no texto.*). O realce da Tarefa 2 como a tarefa que ela considera mais válida é feito em vários momentos, tanto quando descreve o tipo de tarefa que considera mais adequada em geral como quando se refere especificamente à Tarefa 2 (ver exemplos EP032 e EP033, página 70). Com relação à Tarefa 3, a examinanda produziu algumas falas contraditórias sobre a validade da tarefa: em alguns momentos ela atribui certa validade, enquanto que em outras parece indecisa e questiona a validade da tarefa:

EP112	<p>P: Você acha por exemplo que... esse é uma ferramenta adequada de avaliação de leitura?</p> <p>EP112: Então, de leitura... não necessariamente, né, porque aqui você tá avaliando se a pessoa sabe... bom, você vai avaliar se ele leu o texto pelo próprio enunciado da tarefa, pelo o que tá pedindo... é... aí a pessoa tem que usar conceitos que foram ditos no texto... então... é... isso mostra que ela conseguiu ler e entender o texto... mas aí ele tá avaliando escrita também, imagino... porque... às vezes a pessoa pode saber... ler, ler numa boa os textos, consegue interpretar, e na hora de escrever tem uma outra dificuldade né. Que no meu caso é justamente essa, quanto mais tempo você fica sem estudar, as dúvidas aparecem mais freqüentemente sobre... gramática mesmo... né... então acho que aí é um texto que pode de repente a.... pra um corretor, é... se a pessoa que tá fazendo esse teste tem uma dificuldade maior na escrita, pode ficar difícil de passar pro corretor, por exemplo, que ele entendeu o texto...</p>
EP115	<p>P: Mas você acha que serve como uma medida de avaliação?</p> <p>EP115: Então, eu acho que ele avalia, sim. Mas... é... eu acho que o fato de ter que escrever em inglês pode dificultar... dependendo do grau de dificuldade que a pessoa tem pra escrever em inglês... eu acho que... pode atrapalhar também a... que o corretor perceba que ele entendeu o que tava escrito no texto em inglês.</p>

A dúvida sobre se o item avalia adequadamente²⁷ a leitura parece em parte ligada à incerteza sobre se a escrita também é avaliada:

EP118	<p>Se ele tá pedindo, assim como você escreve uma carta no vestibular, ou uma dissertação, que tem uma estrutura própria... um e-mail também tem uma estrutura informal, na maioria das vezes. Principalmente se é pra um amigo, né. É um texto informal. Então possivelmente... acredito que ele esteja esperando isso também.</p>
EP119	<p>Então, tem duas coisas, né. Se é só pra avaliar leitura, eu acho que o estilo do texto não deveria ser considerado. Mas... é... mas... não sei, agora assim tá pedindo um e-mail, pra um amigo, eu não sei se ele tá esperando que eu escreva pra um amigo como EU acho que é... que é um estilo informal... não sei se isso pode contar.</p>
EP120	<p>É... a princípio eu escrevi um e-mail pra... como se eu estivesse escrevendo pra um amigo, estilo totalmente informal. E... ou, acho que em partes, informal... mas eu acho que não fica claro se... se é pra fazer exatamente nesse formato, ou não, mas eu acho que isso também não deve ser... o que tem mais importância nesse tipo de teste.</p>

Essa concepção não parece ligada apenas à Tarefa 3, mas também à Tarefa 2:

EP107	<p>Não, se o teste é pra ser, é pra saber se eu sei ler minimamente em inglês... eu acho que meu estilo de escrita não vai influenciar, não...</p>
EP108	<p>P: Nada?</p> <p>É, a menos que fosse um texto muito cheio de erros, que prejudicasse o próprio entendimento do corretor. Senão acho que não influenciaria, não.</p>

²⁷ Embora termos técnicos da área de avaliação, como “validade” e “autenticidade”, sejam usados no trabalho para análise dos dados, convém ressaltar que, durante as entrevistas, procurei substituí-los por palavras e expressões menos técnicas, como “avaliação adequada” e “avaliar adequadamente” (no caso da validade, por exemplo, usando o princípio de que uma tarefa válida avalia “adequadamente” um determinado construto ou conteúdo).

Nota-se que, na perspectiva dessa examinanda, a integração de leitura e escrita adiciona um fator de insegurança que parece levá-la à conclusão de que, provavelmente, se um examinando consegue demonstrar que compreendeu o texto, o *modo* como ele demonstra isso via escrita não deveria ser relevante. Essa tomada de posição frente a essas tarefas certamente não é aleatória, mas orientada por algum aspecto de seu perfil – seu nível de proficiência na LE, suas experiências com avaliações de língua, os usos que realiza cotidianamente da LE ou outro(s) – que a leva a ter essa percepção sobre a tarefa.

Percepções como essas, demonstradas na entrevista, que apontavam certas tendências, mas que, no entanto, careciam de maior definição, (a ampliação do leque de entrevistados, por exemplo), tornaram o estudo piloto muito valioso no sentido de nortear as reflexões até então feitas apenas no campo teórico. Essa primeira coleta de dados alertou para caminhos possíveis na abordagem do tema, orientou para que elementos novos fossem incorporados e para que os irrelevantes fossem descartados.

Um dos caminhos de entrada no corpus, por exemplo, poderia ser a análise da percepção que a examinanda parece ter sobre alguns fatores como decisivos no seu desempenho em avaliações de leitura, tais como o tempo estipulado para a realização, a ansiedade gerada pela insegurança sobre sua proficiência na LE, o uso de micro-habilidades e estratégias. Lumley & Brown (2004, p. 76), como mencionei na subseção 2.3.2 *Diferentes contextos da autenticidade na avaliação*, sugerem que a administração eficaz do tempo e, conseqüentemente, de estratégias efetivas é determinado, em parte, pelo nível de proficiência lingüística do examinando.

Com relação ao fator “proficiência”, um outro aspecto relacionado pode ser a dificuldade e/ou facilidade percebidas nas tarefas (ver EP034, EP036, EP044 nas páginas 70 e 71) – ou seja, a dificuldade que a examinanda do estudo piloto percebeu nas diferentes tarefas pode ser condicionada por seu nível de proficiência; um outro examinando, de proficiência diferente (maior ou menor) poderia perceber essa questão da complexidade do item de maneira diversa. Esse é um ponto que o estudo piloto demonstrou merecer atenção, pois para Bradshaw (1990, p. 26), por exemplo, uma outra possibilidade a ser considerada com relação à dificuldade ou facilidade que os examinandos percebem numa tarefa é a *complexidade inerente ao item*. Em seus estudos, a dificuldade percebida nos itens não

variava de acordo com o nível de proficiência – mais baixo ou mais alto – dos examinandos, mas era semelhante mesmo para examinandos com diferentes níveis de proficiência lingüística. O *C-test*²⁸ foi apontado, por todos os grupos de proficiência, como o tipo de teste mais complexo, comparado a testes isolados em formato de múltipla escolha sobre conhecimento gramatical e a uma avaliação de leitura com perguntas de respostas curtas. O *formato* do C-test pareceu ser mais complexo, independente do nível de proficiência dos examinandos, embora a autora ressalte que aqueles de proficiência mais básica tenham reagido ainda mais negativamente que os de proficiência mais avançada. Agregar, pois, na coleta efetiva de dados, examinandos com diferentes níveis de proficiência no intuito de investigar se a dificuldade percebida sobre as tarefas variaria significativamente ou não, isto é, se a proficiência influi ou o caso é semelhante ao relatado por Bradshaw (id.), uma questão de *complexidade inerente ao item*, foi um dos pontos elucidados pelo estudo piloto.

Ficou muito clara, mais uma vez, a dimensão relativa da autenticidade nesses dados. A concepção de autenticidade da examinanda do estudo piloto sobre as avaliações estava fortemente baseada em seu próprio uso de LE, e como ela mesma coloca, uma das tarefas poderia se parecer mais com atividades cotidianas de outras pessoas, mas não o eram para ela. Além disso, ela parece atribuir maior validade de face à tarefa que considerou mais autêntica (Tarefa 2, resumo em LM sobre texto em LE), considerando-a uma ferramenta mais acurada de avaliação, e considerou importantes os traços que a caracterizavam para o alcance de seu propósito avaliativo. Isso, a princípio, não é consoante com o estudo de Lewkowicz (2000), cujos dados apontaram para o fato de que a autenticidade não seria uma característica considerada importante por muitos examinandos. O estudo de Lewkowicz (id.) difere em pelo menos dois aspectos do estudo presente: *i*) foi de natureza quantitativa e *ii*) não procurou investigar quais eram as percepções de autenticidade dos examinandos, mas pediu que eles julgassem duas avaliações e, a partir daí, foram tecidas conclusões sobre essas perspectivas. Isso já foi apontado na subseção 2.3.2 como uma limitação daquele estudo, que talvez seja devida, em parte, à própria

²⁸ O *C-test* é um tipo de avaliação semelhante ao cloze, porém apenas a metade de cada palavra sim, palavra não é apagada. (v. Shohamy, 1985a, p. 5)

natureza quantitativa do trabalho, que é mais deficiente na investigação de sutilezas dos dados.

Uma questão que parece ser recorrente, tendo em mente pesquisas já realizadas e os dados aqui apresentados, é a incerteza com que os examinandos se colocam frente a itens integradores no que se refere aos construtos avaliados. A examinanda se mostrou insegura quando inquirida sobre que construto era avaliado ali e se aquele instrumento era adequado para isso. Embora em alguns momentos considerasse que a escrita devia atuar em algum nível na avaliação (quando afirma, por exemplo, que se o texto tivesse muitos erros dificultaria a correção), ela parece não considerar esse procedimento “justo”. Isso seria consoante com resultados de pesquisas sobre exames de proficiência já realizados e citados aqui, como os de Charge & Taylor (1997) e Lumley & Brown (2004).

Um outro elemento que os dados do estudo piloto permitiram iluminar é a relação estreita estabelecida entre validade, autenticidade e facilidade dos itens. No decorrer da entrevista, em muitos momentos, a examinanda tem falas que relacionam a validade do item ao que seria mais fácil para ela (ver exemplos EP027, EP032 e EP033, páginas 69 e 70). Isso é um ponto de partida para algumas indagações: existiria aí uma premissa de que o item mais fácil é mais autêntico devido à proficiência lingüística do examinando (e, no sentido oposto, níveis mais altos de proficiência não encontrariam tantas dificuldades e não estabeleceriam essa ligação facilidade-validade)? Ou a facilidade é percebida porque é identificada em itens mais familiares com o tipo de tarefa feita no dia-a-dia (mais autênticas) e o fato de corresponderem mais a essas atividades faz com que pareçam mais válidas? Mais uma vez, a questão da proficiência lingüística aparece como um fator a ser investigado nesse complexo.

A seguir, em 3.2. *Coleta de dados*, trato de algumas implicações que essa discussão preliminar trouxe para a coleta de dados empreendida subsequente ao estudo piloto, e em 3.2.1. *Perfis dos sujeitos*, descrevo os perfis dos sujeitos que participaram dessa coleta.

3.2. Coleta de dados

A discussão feita na seção anterior atentou para alguns tópicos principais identificados nos dados: *i)* fatores percebidos como influenciadores do desempenho; *ii)* a relatividade aparente da autenticidade, determinada pelo contexto de uso real de LE de cada examinando e a importância dessa característica para os examinandos na constituição da validade do item; *iii)* a percepção dos examinandos sobre que construto é avaliado nos itens integrados, tanto o que eles percebem quanto o que eles acham que os examinadores esperam; e *iv)* o nível de proficiência na percepção da dificuldade/facilidade do item e, conseqüentemente, na atribuição de sua validade ao item dada a relação significativa entre esses elementos.

A partir desses pontos, algumas considerações precisam ser feitas sobre os aspectos que busquei corrigir na coleta de dados subseqüente:

- a testagem do protocolo retrospectivo como talvez uma alternativa às entrevistas semi-estruturadas mostrou que as considerações feitas pela examinanda não foram exatamente otimizadas por esse método, e o mesmo se averiguou no início da coleta de dados subseqüente ao estudo piloto. As considerações feitas muitas vezes pareciam não depender da retrospectão, bastando a leitura, uma análise detalhada dos itens e uma entrevista atenta. Portanto, na coleta de dados, os sujeitos foram solicitados a lerem todas as tarefas atentamente, sem estipulação de tempo mínimo ou máximo para isso, e a entrevista semi-estruturada orientou a explicitação de suas percepções sobre as tarefas.

- mostrou-se claro que o perfil dos sujeitos de pesquisa precisava contemplar diferentes necessidade de uso e vivências com relação à leitura em LE. As percepções demonstradas pela examinanda do estudo piloto demonstram que suas concepções são relativas ao contexto de uso a que está habituada, o que confirma a hipótese inicial de relatividade da autenticidade e pede que outros perfis sejam investigados para que essa hipótese encontre reforço. Identificada essa necessidade, procuramos por sujeitos com diferentes perfis, em diferentes ambientes, e junto a uma escola particular de idiomas que ministra aulas de inglês para um público variado, identificamos diferentes perfis de sujeitos-alunos: que se preparam para exames de proficiência como o TOEFL; que trabalham em multinacionais e precisam se comunicar em inglês com membros de outros

países; que usam o inglês para o lazer, realizando leituras diversas de revistas, sites de internet, livros; que aprendem inglês para o exame do vestibular, entre outros perfis. Alguns dos sujeitos que são descritos na próxima seção provêm desse contexto. Outros foram contatados e convidados a participar da pesquisa devido a seus perfis diferenciados, que poderiam contribuir para a constituição de um corpus heterogêneo, e que pudessem fornecer dados para a discussão da relatividade da autenticidade.

Os procedimentos de coleta de dados foram os mesmos com os oito sujeitos: na primeira etapa da entrevista²⁹, pedi para que falassem sobre sua história de aprendizagem de inglês como LE, seus hábitos e necessidades de leitura em LE, assim como a experiência que tinham com sua avaliação e suas opiniões sobre elas. Na segunda etapa, após terem dedicado cerca de 15 minutos para a leitura e análise individual das Tarefas 1, 2 e 3, coletei percepções e impressões dos examinandos sobre as tarefas, perguntando sobre que aspectos percebiam como mais e/ou menos parecidos com os usos que faziam da LE em suas vidas, e como esses aspectos se configuravam na constituição da tarefa, isto é, se consideravam que a tornavam mais relevante, mais fácil/difícil, mais válida, etc. Logo após a entrevista, os sujeitos tiveram 45 minutos para realizar o teste cloze³⁰ usado como parâmetro de comparação de seus níveis de proficiência em LE.

3.2.1. Caracterização das tarefas

Para investigar como os examinandos percebiam os diferentes graus de integração de habilidades nos itens, as avaliações de leitura que os sujeitos deveriam analisar precisavam apresentar essa característica. Como o objetivo era refletir sobre o conceito de autenticidade presente nessas avaliações, necessária era a utilização de itens que figurassem alguns “usos reais” da leitura em LE e que se prestassem como avaliação de leitura para que os sujeitos pudessem tecer suas considerações sobre a autenticidade que viam (ou não) nesses instrumentos.

Vale lembrar que o objetivo da pesquisa não é propor um tipo de item como mais válido, autêntico ou de maior qualidade, e sim investigar as percepções de

²⁹ Cf. APÊNDICE I.

³⁰ Cf. APÊNDICE II.

examinandos sobre o conceito de autenticidade na avaliação de leitura segundo diferentes tipos de itens, presentes nos mais variados contextos de avaliação. Assim, buscou-se contemplar aqui itens que refletissem diferentes avaliações e que apresentassem uma gradação de integração de macro-habilidades entre si.

Como já discuti na subseção 1.4.2, é arriscado generalizar sobre os usos reais de língua, dada a grande variedade deles e a impossibilidade de abarcá-los todos. Dessa forma, o número de itens utilizados - três - buscou dar conta desses aspectos, além de constituir material razoável para análise feita pelos examinandos - nem muito extenso, nem conciso demais: um item de habilidade independente, tal como de múltipla escolha; um de habilidades integradas que visasse à aproximação com uma tarefa de uso real da língua; e um que representasse um meio-termo, que não integrasse macro-habilidades, mas se revelasse relativamente integrado em termos de proximidade com o uso de língua cotidiano.

Dividirei a caracterização de cada uma das tarefas em subseções, discutindo os princípios que orientaram a seleção de materiais e construção desses itens, brevemente referindo-me ao que a literatura da área de avaliação traz sobre cada um deles.

3.2.1.1. Tarefa 1: Múltipla escolha

O item identificado por *Tarefa 1*³¹ é constituído por uma passagem textual relativamente curta e dez questões de múltipla escolha, e foi retirado de um antigo exame TOEFL, anterior às mudanças que vêm sendo implementadas. Dessas questões, três se concentram na avaliação de conhecimento de vocabulário (questões 2, 6 e 7), uma na de gramática (questão 8), três em averiguação de compreensão mais local dentro do texto (questões 4, 5 e 9), duas em compreensão global (questões 1 e 10) e uma integra conhecimento de vocabulário e compreensão local (questão 3).

Uma das características que torna o item de múltipla escolha vastamente usado em diversas avaliações é a praticidade de sua aplicação. Sua correção é confiável e rápida, não sendo necessário a “mediação” do corretor, como num item que exija produção escrita, por exemplo. Uma de suas maiores desvantagens, no entanto, reside justamente na

³¹ Cf. APÊNDICE III.I.

mediação entre o formulador e o conhecimento que se deseja avaliar: enquanto que em itens que exijam algum tipo de produção – oral ou escrita – é necessário um corretor que emita um juízo sobre o desempenho demonstrado, e esse fator constitua um elemento que muitas vezes pode reduzir a confiabilidade da avaliação, no teste de múltipla escolha o problema recai sobre a formulação do item (cf. Retorta 1996, p. 41). Enquanto a subjetividade do corretor pode resultar num juízo enviesado, a subjetividade do formulador dos testes, quando na elaboração do enunciado ou das alternativas (corretas e distratoras), pode enviesar a própria demonstração de conhecimento pretendida. Seria o caso, por exemplo, da construção de distratores que tornem a questão fácil (ou difícil) demais, ou ambígua, ou algumas vezes impossível de ser solucionada – uma vez que, se estamos falando de textos, devemos nos ater ao cuidado de que não existe uma única leitura possível. É certo falar em limites estruturais e contextuais da interpretação, e que em avaliações de leitura prima-se pela interpretação que esses limites impõem, mas é aí justamente que deve residir todo cuidado e rigor na construção dos itens de múltipla escolha, para não permitir que sejam muito estreitos com relação à compreensão textual almejada, ou amplos demais.

A escolha do item de múltipla escolha como uma das tarefas a serem analisadas se deu, em primeiro lugar, por esse tipo de item representar o que chamamos na seção anterior de item que avalia uma habilidade *independente*: somente a leitura é requerida na sua realização, não sendo requerido desempenho oral ou escrito. Sendo um dos focos da pesquisa a percepção que os examinandos teriam sobre essa característica na constituição da autenticidade de um item avaliativo, o formato de múltipla escolha se conformava a essa necessidade.

Embora não integre macro-habilidades outros tipos de integração são possíveis mesmo num item independente, como procurei demonstrar na seção 3.2: ele pode integrar a avaliação de variados tipos de capacidades e conhecimentos num mesmo item. Podemos falar, então, de itens de múltipla escolha isolados (no sentido de *discrete point*) e integradores (Oller 1979. p. 233-237).

Na Tarefa 1, não se pode dizer que nenhum dos itens seja totalmente isolado; mesmo as questões que se referem à avaliação de vocabulário não o fazem de maneira

descontextualizada, sempre relacionando o vocábulo ao contexto mais imediato que o circunda. Tampouco se pode dizer que esse tipo de item seja integrador no sentido mais atual do termo – que se aproxime do modo como a língua é usada no dia-a-dia, que denote atenção a aspectos sociolingüísticos e discursivos etc.; o próprio formato de múltipla escolha não favorece essa semelhança com o uso de língua real (embora se possa argumentar que, quando em dúvida sobre o significado de uma palavra, automaticamente repassemos mentalmente as alternativas de significado de que dispomos; no entanto, esse seria um processo muito particular e não tão consciente, pelo menos para leitores mais proficientes). No entanto, poderíamos dizer que, de maneira geral, os itens da Tarefa 1 apresentam algum grau de integração em sua formulação, não constituindo itens que se concentrem sobre a aferição de um ponto lingüístico específico.

3.2.1.2. Tarefa 2: Resumo

O item que se propunha como relativamente integrado foi construído como uma proposta de resumo em LM que procurasse sintetizar as idéias principais de um texto em LE. O texto escolhido para constituir a *Tarefa 2* foi retirado de um portal conhecido da internet, numa seção de artigos de vulgarização científica retirados da revista *Psychology Today*.

Scaramucci (1990), buscando refletir sobre o resumo como tarefa de avaliação da compreensão em leitura em LE, aponta como uma de suas características a aplicação de *macroregras* pelo examinando, num processo de redução do texto-base até um nível de significação básica ou global (sua *macroestrutura*), construindo, assim, não uma reprodução do texto original, mas um novo texto. Essas regras aplicadas pelo leitor estariam ligadas a habilidades de diversas naturezas, como conhecimentos lingüísticos, de estrutura textual, valores e crenças, identificação das condições de produção do texto e dos objetivos com que foi escrito e de seus próprios objetivos como leitor. Como se pode ver, várias habilidades estão envolvidas na produção de resumos e, portanto, embora não integre macro-habilidades, a tarefa sugerida prevê a integração de diversos outros tipos de sub-habilidades.

Uma das questões apontadas pela autora como um problema que persiste na validação de resumos como tarefas de avaliação é a caracterização do que seria um “resumo adequado”. A princípio, um que contemple as linhas gerais de significação do texto, dentro dos limites de interpretação impostos pelas pistas contextuais e estruturais fornecidas pelo autor do texto-base, cumpriria essa tarefa. O nível de detalhamento e a relevância dos elementos inseridos no resumo, no entanto, são fatores que não podem ser desconsiderados, e aí se coloca um dilema, pois são fatores que podem variar de acordo com o tipo de texto-base, o propósito de sua leitura, o propósito do resumo etc. Para a autora, mais adequado é especificar o que é um resumo adequado de acordo com cada situação, não sendo congruente firmar o que é um resumo “bom” para todo e qualquer tipo de situação, mas sim para o objetivo a que se presta.

Para a construção da Tarefa 2, assim como para a da Tarefa 3, a seleção dos textos foi demorada e criteriosa: os textos deveriam ser autênticos, retirados de fontes acessíveis aos sujeitos e cujo gênero pudesse figurar num exame de proficiência. Além disso, deveria tratar de um assunto que pudesse chamar a atenção numa prática de leitura comum, não trazendo tópicos muito especializados.

Para o resumo da Tarefa 2, que deve partir de um texto-base de vulgarização científica em que alguns termos técnicos mais específicos são usados de uma forma que seja compreensível para não especialistas da área, um dos pontos importantes seria compreender as relações entre esse processos descritos e suas implicações para a saúde das pessoas. Prender-se ao significado estreito de termos técnicos e fazer disso um impedimento para compreensão dessas relações significaria uma perda em termos de compreensão textual, e, conseqüentemente, também na produção de um resumo adequado.

Embora não exija a produção escrita em LE, não podendo ser considerado, portanto, um item que integre macro-habilidades, a tarefa de resumo se revela integradora em outros sentidos. Como já disse, a escolha desse tipo de item aconteceu justamente por ele, a princípio, parecer representar um tipo de tarefa que se assemelharia com o tipo de atividade comum no uso de LE de muitas pessoas e, em especial, dos sujeitos desta pesquisa. Como afirma Scaramucci (1990, p. 65), o resumo é largamente usado como método para avaliação da compreensão em leitura em LM e LE em contextos escolares, e

daí pressupormos que não deveria constituir um elemento desconhecido para nenhum dos sujeitos.

Um dos perfis usados para a coleta de dados foi o do sujeito do estudo piloto – uma estudante de pós-graduação cuja maior necessidade de uso da LE era a leitura de textos acadêmicos em inglês – e, como nessa etapa identificamos a tarefa do resumo como altamente relevante para esse sujeito, e nos baseando no pressuposto de que o resumo deveria ser um gênero pelo menos experienciado devido ao seu amplo uso em contexto escolar, o resumo foi mantido no papel de tarefa relativamente integrada.

3.2.1.3. Tarefa 3: E-mail

O item de habilidades integradas, *Tarefa 3*, teve seu texto retirado da mesma fonte que o da *Tarefa 2*. A tarefa formulada sobre esse texto prevê a escrita de um e-mail endereçado a um amigo, utilizando-se, nessa produção, elementos retirados do texto. A proposta busca deixar claro que devem ser usados elementos da leitura prévia na escrita pedindo que conselhos práticos fossem dados, interpretando-se os conselhos gerais dados por especialistas no texto lido. Deixa claro também qual o gênero exigido e a quem se dirigia a produção.

A Tarefa 3, a exemplo de tarefas como as do exame CELPE-Bras, por exemplo, propõe a integração das macro-habilidades leitura e escrita e busca simular um tipo de tarefa passível de ser experienciada pelo examinando em seu uso comum da LE. Obviamente, o que é uma tarefa de “uso comum”, como já indiquei anteriormente, é relativo, e a base tomada aqui para a construção da tarefa foi o tipo de situações que exames de proficiência tomam como exemplos de atividades de uso de língua, e ao qual se prestam como instrumento de mensuração de desempenho preditivo aproximado. É importante ressaltar que está fora do escopo dos exames de proficiência prever todos os tipos de situações de uso de língua possíveis, mas sim prover uma amostra de desempenho lingüístico que permita tecer inferências sobre o provável desempenho do examinando para certos tipos de uso, que podem variar muito. A abrangência dos tipos de tarefas dos exames de proficiência devem, portanto, permitir a realização dessas inferências, o que em alguns

casos se dá através de tarefas mais e/ou menos parecidas com o uso de LE que cada um dos examinandos faz, está acostumado ou julga relevante.

Na busca por instrumentos que permitam a realização dessas inferências, examinadores procuram construir tipos de itens que se aproximem do tipo de uso de língua que o contexto do exame sugere – exames dessa natureza são formulados geralmente para a mensuração da proficiência lingüística geral, em todas as macro-habilidades; ou, ainda, prevendo que exames podem produzir impactos no ensino de LE, os itens podem ser construídos almejando que um certo tipo de visão sobre ensino de LE seja adotado, promovendo um efeito retroativo na aprendizagem – assim, não apenas itens que abarquem a avaliação de mais de uma macro-habilidade são construídos, mas que também permitam a avaliação das habilidades do examinando em lidando com diferentes gêneros textuais, interlocutores, propósitos de interação, registros de língua e tantos outros aspectos que examinadores podem escolher contemplar em seus exames.

É segundo esses princípios que a Tarefa 3 foi construída. Além de seu formato integrar as macro-habilidades de leitura e escrita, ela prevê a mobilização de variados conhecimentos e habilidades para sua realização. O examinando deve compreender o texto em níveis globais e locais, assim como ser capaz de localizá-lo numa determinada comunidade discursiva que permita que ele compreenda que os conselhos dados por especialistas no textos têm que ser convertidos numa outra espécie de linguagem, de um registro mais informal, quando transformando-os em conselhos práticos para um amigo. Só esse aspecto já pressupõe um outro tipo de conhecimento, que é a capacidade de diferenciar gêneros textuais como artigos de revistas eletrônicas e e-mails pessoais. Outra habilidade requerida é a capacidade de selecionar as informações relevantes dentro do texto-base e não apenas colá-las ao texto produzido, mas resignificá-las dentro de um outro contexto, que prevê um outro objetivo de leitura, para um outro interlocutor, numa outra comunidade discursiva. Dessa maneira, o item está avaliando também o conhecimento do examinando sobre adequação lingüístico-discursiva. Obviamente, exigindo-se produção em LE, averigua-se o domínio do examinando sobre aspectos como coerência e coesão textuais, léxico, sintaxe e ortografia, adequados ao gênero exigido no enunciado da tarefa.

Embora esteja me referindo apenas ao resumo (Tarefa 2) como a tarefa “relativamente integradora” do estudo, cabe ressaltar que, a rigor, as três tarefas poderiam receber essa alcunha. Como já demonstrei na subseção anterior, mesmo a Tarefa 1 (o “item independente” do estudo) revela-se relativamente integradora no sentido de integrar, em certa medida, a avaliação de mais de um conhecimento e habilidade e não desprover de contexto o elemento avaliado. Da mesma forma, seria complexo propor a Tarefa 3 como o “ideal” da integração, o extremo *integrado* do continuum já referido, como demonstrarei adiante na subseção 3.3.3. Sendo a integração de habilidades imbricada com a autenticidade, e sendo a autenticidade um conceito relativo a tantos fatores tais como as muitas comunidades discursivas, variedades lingüísticas, registros, gêneros, propósitos de uso etc., é complexo propor um item como totalmente “autêntico”. Por extensão, complexo é também propor uma tarefa como totalmente “integradora”. Esse seria mais um conceito virtual, um ideal de que as tarefas integradoras procuram se aproximar, podendo ser “mais” ou “menos” integradoras. Assim, mesmo a Tarefa 3, que se propõe como a mais integradora do estudo, também é apenas relativamente integradora.

A título de estudo, no entanto, e devido à necessidade de conferir uma nomenclatura que descreva plausivelmente as características gerais da tarefa e reflita o que essas características significam na pesquisa, o resumo é identificado aqui como “tarefa relativamente integradora”, entre as três tarefas dos estudos.

Tendo recorrido sobre os materiais sobre os quais se baseou a coleta de dados, enquadrando-os segundo diferentes níveis de integração, consoantes com as definições de integração de habilidades discutidas no Capítulo II, apresento, na seção seguinte, a efetiva coleta de dados.

3.2.2. Procedimentos da coleta de dados

Identificada a necessidade descrita na subseção anterior de contemplar diferentes necessidades de uso e vivências com relação à leitura em LE, procuramos por sujeitos com diferentes perfis, em diferentes ambientes, para reunir uma amostra de sujeitos que pudesse contribuir para a constituição de um corpus heterogêneo, e que pudessem fornecer dados para a discussão da relatividade da autenticidade.

Abarcar sujeitos de diferentes perfis é um dos motivos pelos quais se adotou a análise dos itens, e não sua realização efetiva, como já indiquei anteriormente. Diferentes níveis de proficiência exigiriam itens adequados às diferentes faixas de proficiência abarcadas. Solicitar que sujeitos de diferentes proficiências realizassem uma mesma *Tarefa 1*, por exemplo, poderia enviesar os dados: buscando adequá-la aos níveis mais básicos, níveis mais avançados poderiam considerá-la fácil demais, alterando a percepção que poderiam ter sobre o item com relação ao seu formato, sua integração de habilidades e seu nível de autenticidade. O mesmo poderia se dar no caminho inverso, adaptando-se o item aos níveis mais avançados.

No mesmo sentido, utilizar mais de um tipo de item para cada grau de integração de habilidades - mais de uma Tarefa 1, por exemplo, usando o mesmo formato de múltipla escolha, mas adequando a exigência do item à proficiência lingüística dos sujeitos - introduziria variáveis que poderiam enviesar os dados. Como já disse, uma avaliação deve ser adequada ao nível de proficiência esperado do aluno/sujeito/candidato: textos e itens devem estar coordenados com o que se espera que essa proficiência permita a ele realizar com a LE. Dessa forma, se sujeitos com diferentes proficiências fossem realizar os itens, obrigatória seria a utilização de diferentes textos e itens de avaliação, adequados aos diferentes níveis de proficiência. Esse procedimento ampliaria o número de variáveis presentes, dada a utilização de diferentes instrumentos e a falta de garantia que eles se equivalessem em termos de nível de exigência, grau de dificuldade dos textos e proximidade com usos de língua feitos na vida real.

Considero importante ressaltar, mais uma vez, que o termo *proficiência* é aqui usado como relativo à situação de uso, e não como um “ponto absoluto” que deva ser atingido para que se possa ser considerado capaz de usar a LE³². Assim, as rotulações *básico*, *intermediário* e *avançado* referem-se a gradações de proficiência *dentro do contexto desta pesquisa*. Essas classificações, acompanhadas das especificações dos usos de LE feitos pelos participantes, têm como objetivo fornecer um parâmetro da proficiência dos sujeitos ao usarem a LE, para que comparações possam ser estabelecidas entre os sujeitos e se possa, minimamente, estabelecer um certo padrão de análise dos dados e não podem ser

³² Como já indiquei anteriormente, para uma discussão dessa questão ver Scaramucci (2000).

tomadas como medidas absolutas. Isso iria contra o conceito de proficiência tomado nesse trabalho, que a toma sempre como *relativa* a algo. No caso deste estudo, os sujeitos classificados como *Nível I*, por exemplo, são considerados “básicos” *em relação* aos outros sujeitos que demonstram um nível de proficiência mais avançado, ou seja, essa é uma classificação válida para o universo dessa pesquisa. Não serão feitas considerações sobre as percepções demonstradas por “sujeitos de proficiência básica”, como se esse rótulo pudesse ser aplicado em qualquer situação, mas sim sobre sujeitos que possuem mais e/ou menos proficiência, comparados uns aos outros.

Dessa forma, a partir dos resultados do teste cloze, é possível delinear faixas de proficiência em que se encontram os sujeitos entrevistados, porém sem deixar de atentar para o fato de que o próprio teste cloze constitui uma medida apenas aproximada de proficiência em LE e que seus resultados não são “absolutos”, mas constituem balizas para situar os níveis de proficiência dos sujeitos, comparados uns aos outros. Para efeitos de análise dentro do universo desse estudo, a essas faixas de proficiência dei os rótulos de *Nível I*, para os sujeitos que apresentaram proficiência mais elementar, *Nível II*, para os sujeitos que apresentaram pontuação intermediária na realização do teste cloze, e *Nível III*, para os sujeitos que demonstraram proficiência mais avançada. Esses resultados são fruto da correção do mesmo teste realizada de três maneiras diferentes (v. Scaramucci 1995):

a) *Cloze I*, em que são considerados os acertos da *palavra exata* do texto original, constituindo medida aproximada da proficiência em LE;

b) *Cloze II*, em que são considerados os acertos de *palavras aceitáveis por falantes nativos*; a partir das respostas dadas por falantes nativos de inglês, foi estabelecido um rol de palavras que, mesmo não sendo as palavras presentes no texto original, poderiam ser aceitas como respostas às lacunas do teste. Esse critério de correção também fornece uma medida aproximada de proficiência em LE;

c) *Cloze III*, em que são considerados os acertos de *palavras que preservem o sentido original*, mesmo que não fossem as palavras originais do texto ou as constantes no rol de palavras aceitáveis por falantes nativos. O teste corrigido dessa maneira fornece uma medida aproximada de proficiência em leitura em LE.

O número de acertos, em cada um dos três tipos de correção, foi convertido numa escala de 0 a 10 e comparado aos dos outros sujeitos, de modo a esboçar um quadro com as faixas de proficiência com que essa pesquisa trabalha.

3.2.3. Perfis dos sujeitos

Os perfis dos sujeitos contam com descrições das atividades que realizam, dos usos que relataram fazer da LE em seu cotidiano, um pouco da história de aprendizagem de inglês e das experiências que já vivenciaram com relação a avaliações dessa língua e também com os resultados dos testes cloze aplicados. Eles são indicados por siglas que identificam seu nível de proficiência (I, II ou III) e por uma letra maiúscula que os diferencia dos outros sujeitos da mesma faixa.

A Tabela 1, a seguir, traz os resultados dos testes, em que podem ser comparados os resultados de Cloze I, Cloze II e Cloze III, que representam os acertos alcançados em cada um dos tipos de correção realizados, conforme descrito na seção anterior. Como se pode observar, os sujeitos **I.A**, **I.B** e **I.C** obtiveram resultados aproximados, os escores mais baixos do grupo, tendo sido classificados na faixa de proficiência *Nível I*. Na faixa *Nível II* foram situados os sujeitos **II.A** e **II.B**, e, na faixa *Nível III*, os sujeitos **III.A**, **III.B** e **III.C**, que alcançaram os escores mais altos.

Tabela 1: Resultados dos testes Cloze

Sujeitos	Cloze			Faixa de Proficiência
	Cloze I	Cloze II	Cloze III	
I.A	2,50	3,33	4,16	Nível I
I.B	1,94	3,05	4,44	Nível I
I.C	1,94	3,05	5,55	Nível I
II.A	3,33	5,00	8,33	Nível II
II.B	3,05	5,83	7,50	Nível II
III.A	3,88	6,66	8,88	Nível III
III.B	5,00	6,94	9,44	Nível III
III.C	3,61	6,38	8,88	Nível III

Além dos resultados dos testes cloze, a entrevista prévia à efetiva coleta de dados permitiu caracterizar de maneira mais detalhada as atividades, necessidades e usos

com relação à LE dos sujeitos. Descrevo, a seguir, essas caracterizações, e em seguida apresento a Tabela 2, que condensa esses dados.

A participante **I.A** é a que forneceu os dados do estudo piloto. Como já descrevi anteriormente, ela tem 28 anos de idade e relatou a necessidade de ler textos em inglês-LE durante seu dia-a-dia em seus estudos de pós-graduação. Ela é aluna de mestrado na área de Biológicas numa instituição pública e vive/viveu a realidade de avaliações de leitura em inglês-LE no programa de pós-graduação que frequenta, que exige testes de proficiência no ingresso e no decorrer do curso. A participante ainda relatou ter realizado estudos de inglês em escolas de idiomas por 2 anos. A participante **I.A** demonstrou, através da entrevista, possuir um nível de proficiência que lhe permite ler textos em LE de sua área de atuação e usar as referências daí tiradas na escrita de sua dissertação de mestrado, em LM, e não considera sua proficiência oral suficiente para se comunicar satisfatoriamente. Ainda, no contexto dos outros sujeitos, demonstra possuir um nível de proficiência mais básico, e suas experiências de uso são voltadas para o uso funcional do inglês, não demonstrando usá-lo em outros contextos fora de seus estudos.

A participante **I.B** tem 26 anos, faz especialização lato sensu na área de Biológicas numa instituição privada e trabalha como profissional da área de saúde. Estudou inglês por aproximadamente 3 anos em cursos particulares de idiomas e relatou a necessidade de ler textos em inglês durante seu dia-a-dia em seus estudos de pós-graduação. Como experiência em avaliações da LE, citou os vestibulares e os testes de proficiência de ingresso no programa de pós-graduação. Com relação ao seu nível de proficiência, **I.B** tem um perfil muito parecido com o de **I.A**, demonstrando usar a LE quase que somente para ler textos técnicos da área de seus estudos da pós-graduação, mostrando certa insegurança quanto inquirida sobre sua proficiência no uso da LE para outros fins, como escrever textos (técnicos ou não), traduzir pequenas passagens dos textos que precisa ler, ou conversar com falantes nativos (uma referência comum para leigos sobre o conceito de proficiência oral).

A participante **I.C** tem 26 anos de idade e faz mestrado na área de Biológicas numa instituição pública. **I.C** estudou inglês por cerca de 2 a 3 anos em escolas particulares de idiomas e atualmente usa a LE somente para ler textos acadêmicos que seus estudos de

mestrado exigem. Ela ainda relatou ter que se submeter a avaliações de leitura em inglês-LE no programa de pós-graduação que frequenta, que exige testes de proficiência no ingresso e no decorrer do curso, e a necessidade de preparar-se para o exame TEAP (*Test of English for Academic Purposes*), exigido no ingresso do curso de doutorado que pretende cursar. Os usos relatados da LE são mais voltados para a leitura e tradução de textos técnicos, embora tenha mostrado certa segurança ao referir-se à sua capacidade para ler outros tipos de texto que não técnicos.

O participante **II.A** tem 28 anos de idade e faz mestrado na área de Exatas numa instituição pública. Ele estudou inglês no Ensino Fundamental e Médio e lia textos em inglês durante a graduação, porém sem estudar essa língua formalmente. Durante o mestrado, fez um curso de inglês de algumas semanas num país falante de língua inglesa. **II.A** relatou a necessidade de ler textos em inglês para seus estudos do mestrado e a experiência dos vestibulares de inglês e dos testes de proficiência de ingresso no programa de pós-graduação. Os tipos de usos que relatou ser capaz de fazer com o inglês são: ler satisfatoriamente textos técnicos da sua área e também outros tipos de textos, como artigos de revistas, manuais, jornais, textos da internet. Sua experiência num país falante de língua inglesa demonstra ter aumentado sua proficiência oral, pelos relatos feitos, e o uso atual que faz do inglês, embora majoritariamente voltado para a escrita de sua dissertação de mestrado, demonstra não ser tão funcional quanto o dos participantes do Nível I, por exemplo, visto que se interessa por outras fontes de LE como contato e aprendizagem de inglês, como outros tipos de textos, músicas e filmes.

O participante **II.B** tem 22 anos de idade, faz graduação na área de Humanas de empresas e trabalha no departamento financeiro de uma multinacional. Ele estudou inglês por cerca de 8 anos em escolas particulares de idiomas, e atualmente usa essa língua para ler, escrever e conversar em seu ambiente de trabalho, pois mantém contato direto com pessoas no exterior, por telefone, principalmente, além de ler diariamente jornais em inglês. Suas experiências com avaliações de LE foram o vestibular e os testes de inglês dos processos seletivos em empresas de que participou. Os tipos de usos da LE relatados foram ler, escrever e conversar em inglês em contextos de trabalho e manter contato com várias

fontes da LE, como filmes, músicas, textos de jornais e revistas e comunicar-se por e-mail, e eventualmente pessoalmente ou por telefone, com amigos falantes nativos de inglês.

A participante **III.A** tem 20 anos de idade, faz graduação na área de Exatas numa instituição privada e trabalha no departamento comercial de uma empresa multinacional. Ela estuda inglês há mais de 10 anos, já tendo viajado para países falantes de língua inglesa a passeio e feito um curso de inglês de algumas semanas no exterior. **III.A** relatou a necessidade de ler, escrever e conversar em inglês em seu ambiente de trabalho, devido a seu contato direto com clientes do exterior, por e-mail, telefone e pessoalmente, além do lidar com contratos e documentos escritos em inglês. Sua experiência com avaliações de LE foram os vestibulares que prestou e os processos seletivos de que participou em diferentes empresas. Ela relatou ser capaz de ler textos técnicos de sua área, como artigos, contratos e outros documentos relacionados a seu trabalho, e não-técnicos, como artigos de revistas e da internet. Além disso, seu nível de proficiência permite que se corresponda em inglês por escrito (via e-mail) e se comunique oralmente em inglês por telefone e/ou pessoalmente em contexto de trabalho.

O participante **III.B** tem 25 anos e faz mestrado na área de Exatas numa instituição pública. **III.B** morou 7 anos num país em que o inglês é uma das línguas oficiais, tendo recebido sua educação regular durante esse período em inglês. Ele relatou usar o inglês atualmente para conversar e trocar e-mails com amigos, ler por prazer e para seus estudos. Com relação a avaliações de LE, **III.B** relatou a experiência dos vestibulares e do teste de proficiência do mestrado. Na época da coleta de dados, ele estava se preparando para submeter-se ao TOEFL e para pleitear uma vaga num curso de doutorado no exterior. Os usos que relatou fazer atualmente da LE são manter contato com amigos estrangeiros, por e-mails, telefone e cartas, ler e escrever com grande frequência vários tipos de textos, acadêmicos ou não.

O participante **III.C** tem 27 anos e faz mestrado na área de Humanas numa instituição pública. Ele estudou inglês por cerca de 5 anos em escolas particulares de idiomas e participou de um programa de intercâmbio de 6 meses num país falante de língua inglesa. **III.C** relatou a experiência dos vestibulares de inglês e dos testes de proficiência de ingresso no programa de pós-graduação. Como usos da LE, relatou ler textos acadêmicos,

para seu curso de mestrado, e não-acadêmicos, por prazer, e a necessidade de se comunicar em inglês via e-mail com pessoas de instituições de pesquisa no exterior, sobre assuntos relativos ao seu tópico de estudo.

A Tabela 2 permite visualizar de forma condensada as caracterizações feitas a partir da entrevista realizada:

Tabela 2: Caracterização dos sujeitos

Sujeito	Resultados do Cloze	Formação	Atividades	Aprendizagem de LE	Usos de LE	Experiência com avaliações
I.A, 28, Fem.	Cloze I: 2,50 Cloze II: 3,33 Cloze III: 4,16	PG – Biológicas	Mestrado e trabalho	Cursos de idiomas	Leitura de textos acadêmicos	Vestibular, testes de proficiência da PG
I.B, 26, Fem.	Cloze I: 1,94 Cloze II: 3,05 Cloze III: 4,44	PG – Biológicas	Lato sensu e trabalho	Cursos de idiomas	Leitura de textos acadêmicos	Vestibular, teste de proficiência da PG
I.C, 26, Fem.	Cloze I: 1,94 Cloze II: 3,05 Cloze III: 5,55	PG - Biológicas	Mestrado	Cursos de idiomas	Leitura de textos acadêmicos	Vestibular, testes de proficiência da PG
II.A, 28, Masc.	Cloze I: 3,33 Cloze II: 5,00 Cloze III: 8,33	PG – Exatas	Mestrado	Escola regular no Brasil e curso de idiomas no exterior	Leitura de textos acadêmicos	Vestibular, testes de proficiência da PG
II.B, 22, Masc.	Cloze I: 3,05 Cloze II: 5,83 Cloze III: 7,50	G - Humanas	Faculdade e trabalho	Cursos de idiomas	Ler e escrever documentos e e-mails, conversar com clientes, ler jornais	Vestibular, processos seletivos de emprego
III.A, 20, Fem.	Cloze I: 3,88 Cloze II: 6,66 Cloze III: 8,88	G – Exatas	Faculdade e trabalho	Cursos de idiomas no Brasil e exterior	Ler e escrever documentos e e-mails, conversar com clientes	Vestibular, processos seletivos de emprego
III.B, 25, Masc.	Cloze I: 5,00 Cloze II: 6,94 Cloze III: 9,44	PG – Exatas	Mestrado	Situação de imersão num país falante da LE	Leitura de textos acadêmicos e não-acadêmicos, comunicação com amigos escrita e oralmente	Vestibular, testes de proficiência da PG, preparando-se para o TOEFL
III.C, 28, Masc.	Cloze I: 3,61 Cloze II: 6,38 Cloze III: 8,88	PG - Humanas	Mestrado	Cursos de idiomas	Leitura de textos acadêmicos e não-acadêmicos	Vestibular, testes de proficiência da PG

3.3. Análise dos dados

As percepções dos sujeitos acerca da autenticidade representada pelos diferentes níveis e espécies de integração de habilidades das três tarefas variaram em alguns pontos e coincidiram em outros. Apresento agora alguns desses elementos surgidos,

identificando os tipos de dados segundo categorias temáticas definidas a partir da própria fala dos sujeitos e também impressões específicas de cada uma das tarefas analisadas³³:

3.3.1. Conhecimento de vocabulário e de gramática

O conhecimento de vocabulário e de gramática da LE é apontado como habilidade mais importante na realização de avaliações de leitura em inglês. Esses dois tipos de conhecimento parecem estar muito imbricados, segundo as percepções dos sujeitos:

I.A015	Eu acho que a questão do vocabulário e da gramática é importante , porque senão também você vai ter vocabulário mas vai escrever tudo errado, quando for escrever um texto em inglês... e pode ser que você também tentando, mesmo ler um texto, pode ser que você tenha um entendimento equivocado porque você não sabia um tempo verbal, ou alguma coisa do tipo.
--------	--

I.B018	Pra ficar uma coisa mais correta, né... porque se você não sabe as palavras, não tem como ler . Você pode saber tudo do tema, em português... mas se dá um texto daquele tema numa língua que você não sabe, não adianta porque você não vai entender nada, né?
--------	--

P	Então pra realizar uma prova que você tenha que ler o texto, você acha importante saber o quê?
I.C018	Em primeiro lugar vocabulário , as palavras do texto... não que precise saber todas, muitas vezes dá pra entender o texto sem saber TODAS as palavras... mas... não sabendo quase nenhuma, aí também não dá... mas é aí que entra a gramática ... porque, por exemplo... às vezes você não sabe a palavra, mas você vê que ela termina em ED. Aí se você sabe um pouco de gramática, você sabe que aquilo é um verbo, porque verbo no passado termina em ED... e isso pode te ajudar a entender aquela parte do texto .

II.A017	De leitura? Ah... por exemplo, um bom vocabulário é importante... o aluno ter um entendimento da gramática , tal, saber a... isso é básico, né, que adjetivo vem antes de substantivo, pra não ficar perdendo tempo... porque essas provas geralmente é com tempo, né, rápido.
P	Você acha que o tema do texto interfere?
II.A018	O tema do texto? Acho que interfere. Lá no exame do mestrado interferiu. Depende do que vai focar na área, né... Se o cara for fazer a prova numa área, ele já... já ta mais dentro daquilo. Já conhece o vocabulário usado...

³³ As partes destacadas se referem a elementos que reforçam a identificação desses pontos nas falas dos examinandos; o uso de maiúsculas indica a ênfase feita pelos próprios examinandos, e o uso de reticências indica pausa ou hesitação; Os códigos X.X000 se referem à sigla usada como referência ao sujeito citado e à localização do turno citado na extensão da entrevista; o código P refere-se aos turnos da pesquisadora e entre parênteses estão seus apontamentos.

II.B011	Nas provas, eu buscava assim... desde que passei a ter um nível que eu poderia estar me expressando melhor... eu buscava não só escrever aquilo que tava sendo pedido, mas também desenvolver uma coisa a mais. Pra desenvolver vocabulário, gramática... principalmente gramática , porque se você não tiver isso bem estruturado, você não... não sai do lugar. E você tem que enriquecer seu vocabulário , pra chegar num nível mais elevado.
---------	--

II.B015	Ácho que dependendo do grau de dificuldade, é necessário que você conheça cada palavra, então vocabulário é importante... assim como você ter conhecimento da situação, do contexto que o texto ta apresentando... porque também você pode entender cada palavra e elas não te dizerem...
---------	---

III.A024	Conhecimento de vocabulário.
P	Vocabulário você acha importante...
III.A025	Vocabulário é muito importante. Acho que não tem muito... muito mistério. Se você sabe, se você tem vocabulário... se você sabe usar corretamente os verbos... e tudo mais, e se você tiver calma... que é o mais importante, eu acho... acho que não tem muito problema.

P	Pense agora nos exames de leitura em inglês. O que você acha que é importante pra se sair bem num exame desses?
III.B022	De leitura? Você ter um vocabulário , né. Um vocabulário... que engloba verbos, tempos verbais mínimos... é, um vocabulário, né. Porque sem o vocabulário... dependendo... tem palavras-chaves dependendo do tema... se você perde aquela palavra ali que é crucial... você não pega o... contexto. Então, pelo menos assim, eu sempre parti do princípio que quanto mais vocabulário, melhor.

III.C009	Da prova em si... eu vejo principalmente... pra quem já tem um conhecimento... eu acho que seria o vocabulário... realmente. E também a questão da relação entre as palavras, né... um pouco do universo que a gente usa aqui... então acho é isso, como as palavras se relacionam em inglês... e também pra um caso específico de uma prova, o conhecimento de vocabulário.
----------	--

3.3.2. Integração de habilidades na avaliação

Diferentes percepções com relação à *validade da integração de habilidades na avaliação* apareceram nas falas dos sujeitos nas entrevistas. Alguns sujeitos (sujeitos II.A, II.B, III.A, III.B e III.C) atribuem caráter positivo à integração de leitura e escrita nas avaliações; outros (sujeitos I.A, I.B e I.C) não percebem essas características como traços positivos:

IA027	Eu acho que é mais FÁCIL... não sei se... pra MIM seria mais fácil... ter um texto com questões abertas em inglês e poder responder em português. Porque... aí se eu to respondendo em português, significa que eu consegui ler minimamente o texto e passar pro português , escrever em português, né, o idioma que eu domino.
-------	--

I.A033	Sim eu acho que... porque é um teste de leitura, né? Pra ver se você sabe ler, se você consegue ler um texto em inglês e entender. Então se é pra isso, se é pra ver se você consegue ler, entender aquele texto, eu acho que... é legal que seja um teste pra você responder em português. Se você responder corretamente em português significa que você leu o texto e entendeu.
--------	---

I.A034	É eu acho que... é... esse teste que eu também disse que acho que é mais interessante, que é ler textos em inglês e poder responder em português... porque eu leio textos em inglês... que são referências importantes pro meu trabalho... e eu vou fazer a referência dos mesmos em português... então eu tenho que ter entendido o que tá escrito lá, e aí vou escrever em português, né...
--------	---

I.B018	Acho que na verdade é melhor responder em português , as perguntas... igual na prova que eu fiz, lá na minha especialização... tinha um texto e tinha perguntas pra responder em português , acho que isso deixa você mostrar aquilo que você entendeu daquele texto.
--------	---

P	E essa prova, que você disse que foi bem... como era?
I.C032	Texto em inglês e tinha umas perguntas em português, pra responder em português também.
P	E o que você achou desse... desse formato, dessa prova ser assim, com texto em inglês e perguntas em português e pra responder...?
I.C033	Ah, achei excelente , tanto que eu fui bem. (risos) Porque se fosse pra responder em inglês, eu acho que não teria passado , e teria que estudar mais pra fazer de novo no... no semestre seguinte.

P	E o que você acha das provas que exigem leitura e escrita juntas? Você ler um texto em inglês e responder em inglês.
II.A030	Eu acho bom. Eu acho que você tem que testar não é só a leitura. Você quer avaliar, né. A leitura e a escrita do aluno.
P	A princípio, você seria a favor.
II.A031	Seria.
P	E uma prova que te dá um texto em inglês com perguntas pra você responder em português?
II.A032	Assim... eu acho que não. Deveria ser tudo em inglês . Por mim deveria ser assim, uma prova em inglês, tem que ser tudo em inglês.

P	Você não acha que responder em inglês vai interferir... em você demonstrar o seu entendimento?
II.A037	Não... não. A pergunta não precisa ser tão direta que você pegue no texto e, ah, já tá aqui, já cola... pode ser também uma pergunta que te obrigue a pensar em inglês e passar aquilo pro inglês .

II.B020	O que levaria a estar escrevendo um texto... acredito que é a melhor forma de estar desenvolvendo o inglês...
---------	---

P	Desenvolvendo ou avaliando?
II.B021	Acredito que as duas coisas. Tanto desenvolvendo sua fluência, quanto avaliando. Acho que o teste, esse de... múltipla escolha... restringe as chances de você errar. Porque você já tem a opção correta na sua mão... é só você saber discernir. Agora em teste escrito, não. Você tem que conhecer.
P	Escrito em inglês?
II.B022	É, em inglês. Porque aí você mostra se sabe realmente ou não, se você já tem desenvoltura na língua.

P	Você acha que o jeito que a pessoa escreve, numa prova que exige que ela escreva... o jeito que ela escreve pode interferir na avaliação?
II.B048	Acho que sim... porque aí... por exemplo... a pessoa que vai corrigir pode ter algumas idéias... alguma idéia do que deve ser escrito... e como a pessoa deve escrever... que é diferente do que você passou no seu texto... por exemplo, na prova de alternativa, não tem a subjetividade, é aquilo e pronto. Na outra, aquilo que você escreveu passa por outra pessoa, pelo julgamento dela.

III.A028	Porque uma coisa é... você pegar um texto e traduzir. Outra coisa é você conversar com uma pessoa. Eu acho que você conversar com uma pessoa, um diálogo, é muito mais importante que traduzir um texto. Porque na sua vida inteira, não é só no trabalho... na sua vida inteira você vai precisar do inglês como um meio de manter contato. Tanto em viagem, até ajudar um estrangeiro que está no seu país... tudo mais.
----------	--

III.A045	Eu acho que um texto em inglês, a pergunta teria que ser inglês, lógico. Mas se tem uma pergunta em português, tá testando muito mais a tradução sua. Entendeu? Não o conhecimento seu com a língua inglesa. Mas sim uma tradução sua.
----------	--

P	Entendi. Tá, e você acha que o fato de, por exemplo, essa aqui... a número 3 e a número 2, que essa aqui é a em português e essa é a inglês... mas você acha que o fato de ter leitura e escrita junto... isso dificulta?
III.A083	Com certeza. Dificulta , porque além de você ter o reading, você tem que ter o listening, tem que ter o writing... primeiro, você vai ter que LER... ENTENDER... você vai ter que ouvir você lendo, entendendo... pra depois você escrever. Acho que isso dificulta muito mais, se fosse só ler...

III.B036	(...) eu acho que... eu sempre aprendi inglês não traduzindo, mas pensando em inglês mesmo, entendeu... então eu sou contra traduções de textos. Eu sou a favor de interpretações com perguntas em inglês e respostas em inglês. Mesmo que a sua resposta tenha erros de gramática, mas se você conseguir passar a idéia... é o que vale.
----------	---

III.C014	Sempre em inglês. Tanto a pergunta quanto a resposta.
P	Um texto em inglês com perguntas pra responder em inglês...
III.C015	Sim, e a formulação da pergunta também em inglês. Que também já entra... pra pessoa... como

	diz um professor nosso aqui, o entendimento da pergunta faz parte também da resposta.
--	--

3.3.3. Temas dos textos usados nas tarefas

Os temas dos textos das tarefas parecem atuar como uma variável importante na percepção que os sujeitos constroem sobre as avaliações. Embora sejam analisados nesse estudo os formatos e os níveis de integração das tarefas, os temas dos textos constituíram quase que um foco de atenção primário para os sujeitos. Por essa razão, considero importante incluir os dados que versam sobre esse aspecto e refletir sobre o que podem contribuir para a compreensão da construção das percepções demonstradas sobre as tarefas. Em diferentes graus, variando de sujeito para sujeito, a percepção sobre a tarefa parece ser baseada na proximidade ou preferência que os sujeitos têm com o assunto do texto, e não no formato da tarefa.

P	Você colocou que costuma ler esse tipo de texto em inglês raramente. Aquilo que a gente falou aquela hora...
I.A068	Isso, que normalmente eu leio textos pro mestrado e que são textos específicos da área, e textos científicos, não têm a ver com o assunto desse texto , por exemplo.

P	Ahã. E o que você achou de, você achou que o tema... é um tema que você leria... que você já leu... que você lê... é um tema desconhecido ou...
I.A124	Não, não é um tema desconhecido , mas... nunca li esse tema em inglês, por exemplo... mas eu acho que é um tema bastante difundido.

P	Qual das três avalia melhor a capacidade de leitura?
I.B026	O resumo.
P	Por quê?
I.B027	Porque é mais completo... é um texto mais complexo, um tema mais... tem mais palavras diferentes, dá pra avaliar bem a pessoa... se ela consegue ler...
P	Sei... mas e o formato... não o texto, mas o tipo de tarefa, qual seria melhor pra...?
I.B028	Bom... aí acho que seria a Tarefa 1.

I.C041	Acho que o mais fácil é esse (Tarefa) 2.
P	Por quê?
I.C042	Porque... então, não é exatamente igual, mas estou meio acostumada com esse tipo de linguagem... os textos que eu leio têm esse vocabulário, e esse tema também tem a ver com muita coisa que

	a gente discute na minha área... a obesidade... então é uma coisa que eu acho que eu conheço mais que esse do texto (Tarefa 1), por exemplo.
P	Ta... mas e a tarefa mesmo? Assim... você me falou que o texto é mais fácil... e a tarefa, a tarefa de alternativas, ou a tarefa de resumo, ou a número três, que é o e-mail... sem pensar no tema do texto.
I.C043	Qual é mais fácil?
P	É.
I.C044	Ah, então... então é a número 1.

P	Você acha que o tema do texto interfere?
II.A018	O tema do texto? Acho que interfere. Lá no exame do mestrado interferiu. Depende do que vai focar na área, né... Se o cara for fazer a prova numa área, ele já... já ta mais dentro daquilo. Já conhece o vocabulário usado...

II.A040	Que eu iria melhor. Que eu teria maior facilidade. E o terceiro é o mais complicado, que eu achei. Que é mais pesado, e tal... chato, né? (risos) Chato, e escrever o e-mail, né, pra uma terceira pessoa... o que já dificulta um pouco.
---------	--

P	Embora você não tenha gostado do texto... você gostou mais do primeiro (Tarefa 1).
II.A044	É, achei mais interessante. Prendeu mais a atenção , os outros textos eu queria ler rápido, pra me livrar, sabe? (risos)

II.B015	Acho que dependendo do grau de dificuldade, é necessário que você conheça cada palavra, então vocabulário é importante... assim como você ter conhecimento da situação, do contexto que o texto ta apresentando... porque também você pode entender cada palavra e elas não te dizerem...
P	Conhecimento do tema, você diz?
II.B016	É, conhecimento do tema, da área em específico. Suponha que você vai ler um texto sobre Biomedicina... às vezes nem em português você não vai entender nada.

P	Você costuma ler esse tipo de texto (tarefa 1)?
II.B028	Sim, costumo... gosto muito de ler sobre música, origem dos tipos de música, acho interessante.
P	Qual tarefa te chamou mais a atenção?
II.B029	O texto (da Tarefa) 3... porque conheço várias pessoas que se encaixam nesse negócio de estresse no trabalho.

II.B032	Já o 2... seria o resumo em português... já exigiria que você conhecesse um pouco não só do assunto global, mas é mais específico... porque você vê que fala coisas mais específicas, de linguagem de médicos...
---------	---

P	E o que você achou do texto em si? A temática dele?... É próxima do dia-a-dia...?
III.A052	Com certeza, eu sou a prova disso! (risos)
P	(risos)
III.A053	E com certeza eu acho que as mulheres sofrem muito mais que os homens por causa disso.
P	Você acha que esse tema facilitaria a prova?
III.A054	Facilitaria muito , pra MIM. Pra mim, pelo menos, facilitaria muito.
P	Mas você acha que para outra pessoa... poderia facilitar também, ou... depende de cada um?
III.A055	Ah, acho que depende muito, por exemplo... qualquer pessoa que trabalha o dia inteiro... sabe falar sobre isso . Sabe o que é trabalhar o dia inteiro sob pressão.

III.A066	Ah, eu não achei que tá... esse segundo, eu achei que tá bem longuinho, mas não tá... aquele texto cansativo... aquele texto que chega no final você já não sabe mais o que tá falando... entendeu? Mas, por exemplo, eu acho que o que tem uma história mesmo, mais vida real, mais realidade, é o terceiro (Tarefa 3) . No entanto que você tá lendo, você sabe do que tá falando... você não vai esquecer, porque tem uma continuidade. Tá falando sobre um assunto.
----------	--

P	E o que você achou do tema desse? Do número 1?
III.A068	Ah... sinceramente?... Não achei nada do tema dele .
P	Não achou nada?
III.A069	Não. Não vai interferir nada eu saber do jazz americano ou não.
P	Tá... mas você acha que ele é um tema que dificulta a prova, por ser um tema muito difícil, diferente?...
III.A070	Acho. Não, não que é muito difícil, mas acho que dificulta um pouco por ser um tema DIFERENTE. Não é um tema que você tá acostumado . Geralmente, quando você vai fazer uma prova de vestibular, por exemplo, o que você vai estudar em inglês?... Você vai escrever redações em inglês, tá, mas aí você pega um tema atual , que tá acontecendo na política, ou... na sua vida... financeiramente... geralmente é pra escrever sobre isso que tá acontecendo atualmente. São os pontos críticos.

III.B032	E esse Segundo (Tarefa 2), sobre a obesidade e o fato de fumar... é um texto mais técnico, tem bastante termos científicos aqui, falando de cromossomos, de telômeros, de encurtamentos... um texto biológico científico... mais restrito, né ... seria um nível mais elevado, seria mais um exame... talvez TOEFL, não sei... mas com certeza bem mais... se você fizesse uma comparação... esse texto aqui seria bem mais exigente pra pessoa. (...).
----------	--

III.B033	(...) esse (Tarefa 3) eu achei um intermediário. Interessante o assunto também... só que meio técnico, voltado assim pra parte de biológicas ... tem umas palavras aqui, <i>epinephrine</i> , acredito que seja um hormônio, uma proteína, ou uma... enzima... isso aqui é um texto biológico. Dependendo do público que vai ter acesso a um texto desses pra ter conselhos, pra dar conselhos... acho que seria meio restrito , num tipo de exame de vestibular... se fosse pra um exame, por exemplo, pra área de saúde... pra um médico... acho que talvez os exames devessem
----------	---

	ser mais voltados pra cada área de atuação, porque cada profissional vai ler uma coisa mais voltada pra sua área. (...)
--	--

P	Então, mesmo estando acostumado, a (Tarefa) 3 seria a mais difícil?
III.B045	É, porque exige mais de você. Exige que leia ,entenda, depois reescreva sobre o tema... que não é um tema assim, comum... já a (Tarefa) 1 é um tema mais facinho de fazer.
P	Mas tirando a parte do tema... suponha que a Tarefa 1, múltipla escolha, fosse com o texto da (Tarefa) 3...
III.B046	Hã... o que tem?
P	Mesmo assim a (Tarefa) 1 seria a mais fácil então? Porque você disse que o problema é o texto... mas vamos pensar na parte da tarefa mesmo. A tarefa de escrever um e-mail e a tarefa de múltipla escolha.
III.B047	Ah, tá... então, acho que ficaria mais difícil, mas ainda não tanto quanto o e-mail. (...).

P	E você acha que... por estar acostumado a fazer isso... fazer isso numa prova seria mais fácil para você?
III.B053	Então, é aquilo que eu disse... estou acostumado e acho que não seria complicado fazer, mas com esse texto aqui (da Tarefa 2), acho que seria mais difícil. Porque é aquilo que eu já falei, dos três, esse é o tema mais complexo, tem palavras mais difíceis...
P	Tá, mas imagina que não fosse um texto que você considere muito difícil então. Imagine que fosse com um outro texto, assim, intermediário, sem essas palavras difíceis.
III.B054	Bom... então, sim, seria mais fácil. Principalmente se fosse com um tema da minha área, que eu já conheço os termos, as expressões, já teria mais jeito pra fazer. Mas mesmo que não fosse da minha área, se fosse um tema assim... conhecido, difundido... mais banal mesmo, mais corriqueiro... aí seria fácil fazer o resumo.

P	Qual tipo de tarefa... qual formato... numa avaliação... seria mais parecido com o que você costuma fazer com o inglês no seu dia-a-dia?
III.C016	Eu acho que um formato interessante seria pra que eu desenvolvesse um texto em inglês em que eu, por exemplo... me situasse dentro do meu campo de trabalho. Ou... onde... em que área, em que subáreas, está a minha pesquisa... os resultados que eu tenho obtido... me colocaria dentro dessa área e falaria sobre isso.

3.3.4. Tarefa 1

Diferentes percepções com relação à Tarefa 1 foram demonstradas. O tema do texto da Tarefa 1 dividiu opiniões: enquanto que o tema não parece interferir nas percepções demonstradas por I.A, II.A e II.B, nas falas de I.B e I.C o tema parece chamar mais a atenção do que o próprio formato da tarefa. Para estes últimos, a preocupação com

entender o assunto do texto parece um fator importante. Uma explicação possível poderia ser o tipo de leitura que esses sujeitos empreendem cotidianamente: talvez por possuírem proficiência mais básica, I.B e I.C estejam acostumados a realizar leituras de entendimento mais global, sem se ater a detalhes; assim, a compreensão da idéia geral do texto para esses sujeitos é mais importante que para os outros. Já II.A e II.B não consideram os textos usados adequados, III.B por julgá-lo muito simples e III.A, muito fora da sua própria realidade de leitura.

Com relação à autenticidade da tarefa, nenhum dos sujeitos considera o formato da Tarefa 1 parecido com os usos que realizam da LE. Todos associam a tarefa de múltipla escolha não à prática de uso “real” da língua, em situações autênticas de comunicação, mas a práticas escolares institucionalizadas, como exercícios, estudos e avaliações.

Um outro fator interessante foi a questão das alternativas do formato da Tarefa. I.B e I.C percebem as alternativas como uma fonte de ajuda na compreensão do texto e como orientadoras no sentido de realizar a tarefa – como se elas fornecessem uma “pista” para a resposta correta. Já I.A e II.A vêem a múltipla escolha como um formato que pode confundir o examinando, e II.A levanta um ponto interessante, que é o fato de que alguns itens da Tarefa 1 se concentraram sobre aspectos da interpretação do texto que ele não havia considerado quando lendo o texto pela primeira vez. Isso faz pensar sobre a validade desses itens, uma vez que podem estar se referindo a aspectos habitualmente não contemplados na leitura comum. Em sentido semelhante, II.B III.A, II.B e III.C percebem as alternativas com um caráter cerceador da interpretação do texto e da demonstração dessa interpretação em situação de avaliação, pois elas impediriam o examinando de colocar o que realmente entendeu do texto, estando sujeito à adequação às alternativas – que, além disso, os impediriam também de demonstrar outros aspectos de seu conhecimento da LE através da escrita.

P	Você acha que a [Tarefa] 1, por exemplo, é uma atividade totalmente fora do seu... da sua vida real... parcialmente...
I.A053	Não, totalmente fora. Não faço nada desse tipo.

I.A066	Isso. Linha 9, eu acho que ele quis dizer que é um de seus empregos mais regulares, mesmo, né, mais freqüentes, mais regulares, tal... mas na hora de escolher a palavra que poderia substituir,
--------	--

	palavra sinônima, no caso... eu excluí duas, as alternativas c e d... mas como eu não sabia, por exemplo, o que significa a alternativa b, eu posso ter errado por esse motivo.
P	Então na verdade seu problema não foi com a palavra do texto, foi com as alternativas.
I.A067	Sim, sim. Mas eu acho que o sentido deu pra compreender.

I.A089	Não, acho que existem coisas que... acho que é importante por exemplo interpretar direito o texto, mas... é... por exemplo, procurar palavras sinônimas.. não sei... não sei se sempre é uma boa... não sei se SEMPRE isso vai medir que você entendeu o que tá escrito no texto... por exemplo, quando você tem duas alternativas... você conseguiu ler o texto e sabe o que ele tá dizendo... agora você tem duas alternativas e... você não tem certeza das palavras, por exemplo... se são sinônimas... não significa que você não... entendeu o que tá escrito no texto, né...
--------	--

P	Qual das três tarefas avalia melhor a capacidade de leitura?
I.B026	O resumo.
P	Por quê?
I.B027	Acho que é mais completo... é um texto mais complexo, tem palavras mais complexas...
P	Tá... mas e com relação ao... ao tipo de tarefa mesmo... independente do texto, mas o formato... de múltipla escolha, de resumo ou de email?
I.B028	Bom... aí acho que seria a Tarefa 1.
P	Por quê?
I.B029	Por causa dos... das questões de múltipla escolha. Dá pra entender melhor o texto através das perguntas.

P	A (Tarefa) 3? Nada parecido? E a (Tarefa)1?
I.B032	Não.
P	O fato de a Tarefa 2 ser mais parecido com o que você faz no dia-a-dia, tornaria ela mais fácil pra você fazer?
I.B033	Eu acho que não. O mais fácil acho que é o 1 (Tarefa 1).
P	Não pensando no texto, mas no formato.
I.B034	Acho que o 1 (Tarefa 1).

P	Qual o traço importante que o corretor ia analisar... o que é mais importante?
I.B36	No texto 1.. Se ele tem uma compreensão do texto.

P	A (Tarefa) 1 se parece com alguma coisa que você faz no seu dia-a-dia com o inglês?
I.C045	Não, não parece, não costumo fazer muito provas assim, é mais dissertativa.
P	E mesmo não fazendo nada parecido, ela ainda é a mais fácil?

I.C046	Então, se for ver, o resumo é o que eu mais faço. Mas, numa prova, não acho que seria mais fácil, acho que o mais fácil é esse, o de xisinho (sic) (Tarefa 1). Porque às vezes você não entendeu o texto assim... assim tão profundamente pra escrever uma resposta, mesmo que seja um resumo. Já com as alternativas, você lê, aquilo te dá uma idéia, volta pro texto, lê de novo... e daí você pode ir eliminando... claro que às vezes as alternativas são tão parecidas que confundem.. aí depende também...
--------	---

II.A039	O tema 1... é o que é o mais agradável, que é... pra ler assim... mas na questão do teste ali... acho que eu teria que pensar em alguns itens ali pra responder.
P	Como assim? Você teria mais dificuldade?
II.A040	Não sei, achei o texto legal e entendi, não achei difícil... mas as perguntas... perguntando o que é tal palavra, o que significa... são coisas que eu nem tinha pensado a hora que tava lendo, por isso teria que pensar mais pra responder.

P	Você acha que a Tarefa 1 reflete algum uso que você faz do inglês no seu dia-a-dia?... Em algum momento, quando você está lendo... em inglês...
II.A056	Bom, quando eu vou fazer uma prova, né, aí tem esse tipo de exercício.
P	E fora isso? Fora situação de prova?
II.A057	Bom, só quando eu vou estudar...

II.B030	Acho que [a Tarefa 1] é mais comum, mas eu não gosto muito. Porque, primeiro, que você pode entender o texto, por exemplo... se você tivesse lendo o texto na sua casa, ler e entender tudo... só que aí você tem que fazer a prova, e tem que escolher entre a, b,c e d... e nenhuma dessas alternativas é aquilo que você entendeu... ou até pode ser, mas tá escrito de uma forma diferente, pra confundir.
---------	---

II.B037	[A Tarefa 1] Deixa a desejar.
P	Por quê?
II.B038	Restringe demais... não deixa você ampliar, sabe... você fica preso naquelas alternativas, e não deixa você demonstrar outras coisas... demonstrar seu vocabulário, seu desempenho... seu entendimento do texto.

II.B044	O texto 1... mais raramente ainda... algumas vezes tive prova na faculdade usando múltipla escolha, mas hoje em dia... não tem nada a ver com o que eu faço. Seria mais o email mesmo.
P	Múltipla escolha seria mais em situação de avaliação mesmo.
II.B045	É. Só prova, não de uso prático, assim...

II.B012	Eu acho que é um teste mais primário... mais primário, que eu digo, porque ele tá de uma possibilidade de um chute... tem mecanismos pra você... enganar um pouco você, o que inibe um pouco o chute... mas de qualquer maneira, pode contribuir pra que você... se você tem alguma
---------	--

	dúvida, você pode julgar as respostas lá, pra ver qual faz mais sentido. Então... de alguma forma ele te dá alguma idéia da resposta , mais do que uma questão aberta, que você vai colocar aquilo que você sabe mesmo. Não tem nenhum caminho já indicado sobre as respostas. Então eu acho que esses testes de múltipla escolha, eles são mais por questões de praticidade da correção... e rapidez. Mas eu acho que em termos de avaliação... ele é um pouco menos eficiente . Não sei se é eficiente porque não estudei sobre isso, mas o que... eu acho é que ele testa menos . É minha opinião.
--	---

III.A039	Se fosse em, se fosse... dissertativa, acho que seria muito mais fácil. Porque você escreve o que você tá pensando, que você entendeu. Agora múltipla escolha, não. Você entendeu, sabe do que tá falando, mas tem quatro alternativas que pra você são iguais . Tão falando da mesma coisa, mas são totalmente diferentes.
----------	--

III.A057	Eu creio que sim. Mas por exemplo, esse primeiro texto aqui, olha (Tarefa 1)... é com certeza o exemplo de uma prova de inglês. Assim que a gente é cobrado . Mas na verdade, na vida real, é só como essa terceira tarefa que a gente é cobrado.
----------	--

P	E esse aqui, o primeiro (Tarefa 1)?... O tamanho dele...
III.A067	Ah... o ideal. Pra uma prova de vestibular, então, é o ideal. Pequeno, simples...
P	E o que você achou do tema desse? Do número 1?
III.A068	Ah... sinceramente?... Não achei nada do tema dele.
P	Não achou nada?
III.A069	Não. Não vai interferir nada eu saber do jazz americano ou não.
P	Tá... mas você acha que ele é um tema que dificulta a prova, por ser um tema muito difícil, diferente?...
III.A070	Acho. Não, não que é muito difícil, mas acho que dificulta um pouco por ser um tema DIFERENTE. Não é um tema que você tá acostumado. Geralmente, quando você vai fazer uma prova de vestibular, por exemplo, o que você vai estudar em inglês?... Você vai escrever redações em inglês, tá, mas aí você pega um tema atual , que tá acontecendo na política, ou... na sua vida... financeiramente... geralmente é pra escrever sobre isso que tá acontecendo atualmente. São os pontos críticos.

III.A082	Na número 1? Interpretação do texto. Acho que se você souber gramática, vocabulário... aqui você resolve tudo, e isso não significa que você fala inglês ou não.
----------	---

P	Então qual dessas tarefas... fazer múltipla escolha (Tarefa 1), fazer um resumo (Tarefa 2), ou escrever um e-mail (Tarefa 3)... é mais parecido com o que você faz com o inglês no seu dia-a-dia?
III.B030	O e-mail (Tarefa 3), com certeza. Porque aí tem que ler um... um texto, que eu acho que é de internet... eu leio muito na internet... mandar um e-mail... que eu também mando e-mails pros meus amigos que não são daqui, então tem que ser em inglês... o resumo eu faço bastante também, mas acho que menos, e mais por obrigação também. Já a múltipla escolha eu faço quando tenho

	alguma prova, tenho que estudar, me preparar... aí pego provas, exercícios, e me preparo fazendo antes.
--	--

P	Qual dos três você considera que é um instrumento adequado de avaliação de leitura?
III.B037	Pra qualquer nível de pessoa? Porque você fala leitura assim... de uma pessoa que tá na oitava série, um aluno universitário...?
P	Num exame de proficiência.
III.B038	Tipo TOEFL?
P	É, desse tipo.
III.B039	Esse daqui (Tarefa 1), de forma alguma.
P	O número 1?
III.B040	É.
P	Por quê?
III.B041	Porque... esse é muito básico . Acho que ele é muito direto, muito fácil , as questões... acho que uma questão já responde a outra... e o vocabulário dele é simples, é direto... é muito fácil.

III.C017	Esse primeiro aqui (Tarefa 1) é o mais comum... marcar uma opção... pelo que percebi rapidamente, envolve um pouco do que você entendeu do texto.. mas o mais importante é que ele pega aqui... o que eu acho que ele enfatiza mais, eu acho que é a estrutura do texto.. de palavras assim... e se você entendeu o sentido com que ela palavra foi usada. Então mais um conhecimento de inglês do que de... como fala? Quando você... faz uma interpretação de texto. Aqui a maioria são coisas mais isoladas... só aqui nessas perguntas, “a passagem é sobre isso, ou sobre aquilo”... mas é uma interpretação mais geral
----------	--

P	Me fala um pouco sobre os três em termos do que... se são parecidos com alguma coisa que você faz com inglês no seu dia-a-dia... ou se não são parecidos...
III.C019	Acho que o segundo (Tarefa 2) é mais parecido com o que eu faço normalmente. Eu tenho textos para ler em inglês e preciso às vezes elaborar um outro texto, em português, a partir deles... pra colocar na minha dissertação, ou em algum trabalho... o terceiro (Tarefa 3) eu faço de vez em quando também, mas é menos. Às vezes tenho que mandar um e-mail pra algum lugar, com alguma dúvida, e às vezes são instituições no exterior. Aí tenho que escrever em inglês o e-mail. Já o primeiro (Tarefa 1), praticamente não faço isso, acho que a última prova de inglês que fiz de alternativa foi no vestibular, depois só dissertativa, que eu me lembre.

III.C020	Eu achei o primeiro mais fácil. Mas a sua pergunta é... qual dos três seria mais adequado pra uma avaliação, né?
P	É.
III.C021	Eu acho que seria o terceiro (Tarefa 3).
P	Por quê?
III.C022	Porque quando ele vai desenvolvendo o texto... por exemplo, grande parte dessas perguntas que tem no primeiro (Tarefa 1)... tipo, o que significa essa palavra no texto... já tem embutido no

3.3.5. Tarefa 2

Diferentes percepções foram demonstradas com relação à Tarefa 2. Com relação à autenticidade da tarefa, os sujeitos I.A, I.B, I.C, II.A, III.B E III.C a perceberam como parecida com os usos que realizam da LE. Analisando os perfis desses sujeitos, todos estudantes de pós-graduação, fica fácil perceber essa coincidência, visto que a elaboração de textos em português baseada na leitura e interpretação de textos em inglês, a partir das falas desses sujeitos, demonstra ser uma atividade rotineira para esse tipo de perfil. É interessante notar, no entanto, que mesmo entre esses sujeitos as percepções variam. Para I.A, de proficiência mais básica, por exemplo, por se configurar como uma tarefa próxima do uso que faz de inglês, a Tarefa 2 seria a melhor avaliação de leitura entre as três; já para II.A, da faixa de proficiência intermediária, exatamente por estar acostumado a realizar esse tipo de atividade, a tarefa não seria um bom instrumento de avaliação – o sujeito considerou a Tarefa 3 como o melhor instrumento para esse fim. Os sujeitos II.B e III.A relataram que os usos mais rotineiros que realizam com a LE são em ambiente de trabalho, em que precisam conversar e escrever em inglês, fugindo, portanto, do formato da Tarefa 2.

No que se refere a outros aspectos, essa foi a tarefa sobre a qual as percepções mostraram-se mais díspares. Seguindo a linha da percepção demonstrada sobre a Tarefa 1, I.A considerou a Tarefa 2 a mais apropriada para o examinando mostrar seu real conhecimento da LE, não sendo confundido ou ficando preso às interpretações sugeridas nas alternativas da tarefa de múltipla escolha. I.B e I.C, que haviam considerado a múltipla escolha como uma tarefa fácil, perceberam maior dificuldade para a realização da Tarefa 2, embora parecida com os usos que costumem fazer da LE. Para eles, a escrita, mesmo em LM, adicionaria um elemento complicador a mais, além da retirada do elemento facilitador da Tarefa 1, as alternativas. Mais uma vez, talvez isso se explique pelo nível de proficiência mais básico, desses sujeitos, que talvez se sintam mais inseguros sobre a interpretação que são capazes de construir de um texto em LE. No entanto, para III.A, de proficiência avançada, a escrita em LM também seria um elemento de dificuldade – III.A, no entanto,

atribui essa dificuldade ao tipo de linguagem que acredita ser necessário para a realização da tarefa, mais formal, por se tratar de um resumo.

Com relação ainda ao produto da tarefa, I.B, II.A, III.A E III.B percebem a elaboração de um resumo como uma tarefa de *tradução* da LE para a LM. II.A, embora não mencione o termo *tradução* em sua fala, deixa claro que o produto da Tarefa 2 deve ser um texto fortemente calcado no texto original em inglês, e inclusive menciona isso como um fator facilitador da tarefa, em comparação com a Tarefa 3. Já II.B percebe o caráter dessa produção de texto como um texto outro que não o original, e não apenas tradução ou paráfrase, mas um texto de outro gênero, uma síntese.

No que se refere à correção desse produto, II.B e III.A acreditam que a escrita, e não somente a leitura, é avaliada nessa tarefa. Já I.A e III.C mencionaram que a forma como o texto foi escrito não deveria ser avaliada, apenas seu conteúdo, ou seja, a interpretação do texto demonstrada. Esse dado traz um aspecto já discutido no Capítulo II: o aspecto da contaminação da avaliação (*muddied measurement*, Weir 1990). Quando esses sujeitos mencionam a questão, parecem preocupados com o fato de a escrita interferir na demonstração de desempenho na leitura. Podemos ver esse aspecto também na fala de II.B, que julga a Tarefa 2, mesmo menos parecida com os usos que faz da LE em sua vida, a mais adequada como avaliação de leitura, por considerar que, nela, elementos próprios da habilidade da escrita, exigidos na Tarefa 3, não interfeririam na avaliação. A fala dos sujeitos I.B e I.C, que não percebem a integração de leitura e escrita como um fator positivo para a avaliação, também parece remeter à essa questão da avaliação contaminada.

Fica claro que esse aspecto da contaminação, um dos mais controversos quando se tratando de integração de habilidades na avaliação, representa uma fonte de questionamento não apenas para teóricos e professores, mas segundo a perspectiva dos examinando também é um fator de preocupação, principalmente no que se refere a ser “bem” avaliado. A preocupação maior dos examinandos parece ser a de sofrer algum tipo de injustiça, seja por não poder mostrar toda sua capacidade no que concerne o conhecimento da LE, seja por não ser capaz de mostrar o pouco que é capaz por culpa da formulação do teste.

I.A045	Hum... é... eu acho que... a segunda atividade (Tarefa 2), que é aquela do resumo, se parece mais
--------	---

	com um exame de leitura
P	Por quê?
I.A046	Porque eu acho que, bom... porque eu acho que... se o candidato entendeu o texto, de um modo geral... ele tem condições de no seu próprio idioma poder dizer o que entendeu a respeito.

P	Não, sim. Sua opinião. Então você acha... qual das três atividades você acha que avalia melhor?
I.A049	Eu acho que é a segunda.

I.A057	Eu acho interessante. Se eu tivesse que classificar, dentre as tarefas, assim... a que eu escolheria pra eu fazer, que se parecesse mais com meu cotidiano , seria a dois, como eu já falei, depois o e-mail e por último a tarefa número um, que é a de perguntas-teste.
--------	--

I.A105	Então, é que... é... eu acho que assim, o fato de escrever em português facilita bastante... é bem mais fácil... mas, aí tem as outras questões, né... que como esse texto era mais longo que o primeiro, tinha palavras que eu não sabia, não tinha o dicionário... então tem essas outras coisas que dificultaram um pouco mais... mas, pensando em... só na atividade... sem pensar no texto em si, nas dificuldades que eu tive com vocabulário... eu acho que esse tipo de atividade é mais fácil por poder escrever em português aquilo que eu entendi do texto. E aí eu acho que o fato também de fazer isso... é... com uma certa frequência ajuda.
P	E você acha que uma pessoa que fosse corrigir esse texto... no que ela vai mais prestar atenção? Na sua produção aqui.
I.A106	Hum, eu imagino que ela vá... querer saber mesmo se eu entendi o texto, aí ela vai procurar as idéias principais do texto em inglês no meu resumo em português.
P	Ela vai tentar achar se você conseguiu... reproduzir, parafrasear... ou reinterpretar... o que você acha?
I.A107	É, seu eu consegui.. interpretar...acho... acho que é isso.
P	Mais alguma coisa?... Você acha que em termos de escrita, sua escrita vai pesar alguma coisa?...
I.A108	Não, se o teste é pra ser, é pra saber se eu sei ler minimamente em inglês... eu acho que meu estilo de escrita não vai influenciar, não...
P	Nada?
I.A109	É, a menos que fosse um texto muito cheio de erros, que prejudicasse o próprio entendimento do corretor. Senão acho que não influenciaria, não.

P	Qual das três provas reflete o uso do dia-a-dia que você faz do inglês?
I.B030	A (Tarefa) 2.
P	Você faz coisa parecida com isso?
I.B031	É, porque... além de encontrar artigos, eu encontro resumos, né... então eu leio bastante resumos também.

P	Qual o traço importante que o corretor ia analisar... o que é mais importante?
I.B036	No texto (Tarefa) 1.. Se ele tem uma compreensão do texto.
P	No (Tarefa) 2?
I.B037	No (Tarefa) 2... é se ele sabe fazer uma tradução... claro que ele não vai traduzir ao pé da letra, mas vai traduzir...

P	Entre as Tarefas 2 e 3, qual é a mais fácil?
I.B042	A (Tarefa) 2. A (Tarefa) 3, por ter leitura e escrita, é a mais difícil.

I.C041	Acho que o mais fácil é esse (Tarefa) 2.
P	Por quê?
I.C042	Porque... então, não é exatamente igual, mas estou meio acostumada com esse tipo de linguagem... os textos que eu leio têm esse vocabulário, e esse tema também tem a ver com muita coisa que a gente discute na minha área... a obesidade... então é uma coisa que eu acho que eu conheço mais que esse do texto (Tarefa) 1, por exemplo.
P	Ta... mas e a tarefa mesmo? Assim... você me falou que o texto é mais fácil... e a tarefa, a tarefa de alternativas, ou a tarefa de resumo, ou a número três, que é o e-mail... sem pensar no tema do texto.
I.C043	Qual é mais fácil?
P	É.
I.C044	Ah, então... então é a número 1.

I.C046	Então, se for ver, o resumo é o que eu mais faço. Mas, numa prova, não acho que seria mais fácil, acho que o mais fácil é esse, o de xisinho (sic) (Tarefa 1). Porque às vezes você não entendeu o texto assim... assim tão profundamente pra escrever uma resposta, mesmo que seja um resumo. Já com as alternativas, você lê, aquilo te dá uma idéia, volta pro texto, lê de novo... e daí você pode ir eliminando... claro que às vezes as alternativas são tão parecidas que confundem.. aí depende também...
--------	---

P	E você acha que, pra avaliar leitura, ela (Tarefa 3) é uma... prova, uma... ferramenta... boa? Ela avalia bem a leitura que a pessoa faz em inglês?
I.C050	Eu acho que avalia, porque pra escrever o e-mail, a pessoa tem que ter entendido bem o texto, pelo que ta pedindo aqui no enunciado. Mas não é SÓ a leitura, tem a escrita também. Aí, se você tem uma dificuldade maior pra escrever, como é o meu caso, você sai prejudicado, entendeu? Então tem leitura aí no meio, mas não é só isso. Por isso que eu acho que se é pra avaliar leitura, a melhor é a (Tarefa) 2, que exclui essa parte da escrita.

II.A039	O tema 1... é o que é o mais agradável, que é... pra ler assim... mas na questão do teste ali... acho que eu teria que pensar em alguns itens ali pra responder. Acho que no resumo, mesmo não tendo gostado tanto do texto... acho que não teria um entendimento tão bom quanto do primeiro... acho que eu me sairia melhor no resumo.
---------	--

P	E o (Tarefa) 2 seria o mais fácil...
II.A043	É, pra resolver, sim...

P	Você acha que algum desses três testes se parece com alguma atividade que você faz no seu dia-dia? Com o inglês?
II.A050	O segundo.
P	O segundo?
II.A051	É.
P	Você geralmente faz resumos em português de textos em inglês.
II.A052	Faço. É. Tenho um texto em inglês e preciso passar pro português e tal.
P	Mesmo assim você acha que isso não serviria... como uma avaliação... não seria tão adequada?...
II.A053	Então, por eu estar tão acostumado, eu acho que eu... to meio puxando pra mim, né. Por estar tão acostumado, eu acho que não avalia tanto quanto o outro que... além de... além de que teria mais trabalho, vamos dizer, teria que me empenhar mais, pensar mais pra fazer.

P	E essa aqui... a 2. O que você acha que você tem que demonstrar no seu resumo que vai ser um traço valorizado?
II.A059	A idéia principal do texto. Não pode perder a.... tem que passar isso no seu resumo.

II.A061	(...) porque a segunda ainda você pode fazer um jogo de montagem ali, eu acho... pode encaixar... (...)
P	Você falou isso, que aqui na 2 dá pra encaixar... o que você quer dizer com isso? Esse aqui possibilita um uso de estratégias...
II.A063	É, diferente desse 3. Porque um resumo... você vê, aqui você estaria escrevendo pra uma outra pessoa... uma coisa nova, você não... não pode deixar, como que posso dizer...
P	Parafrasear? Você não pode parafrasear?
II.A064	Nesse terceiro? Acho que muda um pouco, né?... A dificuldade é maior.
P	E no 2?...
II.A065	Nesse, eu faria... tem o texto, eu separaria primeiro a idéia central, e a partir dela ia montando o texto. Seguindo a ordem, teria que manter a seqüência também, de cada parágrafo. Acho que você não se perderia na escrita tanto quanto no 3.

II.B032	Já o 2... seria o resumo em português... já exigiria que você conhecesse um pouco não só do assunto global, mas é mais específico... porque você vê que fala coisas mais específicas, de linguagem de médicos...
---------	---

P	E a Tarefa 2, serve como avaliação de leitura?
II.B039	Sim... esse seria mais adequado pelo fato de abranger outras coisas, outro tempos verbais... e você adquire maior vocabulário.

P	A Tarefa 2 se aproxima do seu dia-a-dia? É o tipo de coisa que você faria com o inglês no seu dia-a-dia?
II.B042	Não... acho que raramente .

P	O que o corretor vai valorizar na (Tarefa) 2?
II.B046	Acho que no resumo... a capacidade fazer uma boa síntese do texto... ser capaz de passar o que você entendeu com outras palavras... porque algumas pessoas pensam que resumo é sinônimo de cópia, e não é. Acho que no (Tarefa) 2 então é a capacidade de passar uma idéia semelhante do entendimento que você buscou no texto.

P	Para avaliar leitura , a (Tarefa) 3 é adequada?
II.B049	É, mas a (Tarefa) 2 é melhor.
P	Por quê?
II.B050	Porque acho que ter que escrever um resumo é melhor pra ver se a pessoa entendeu o texto do que pedir pra ela escrever um outro texto sugerindo coisas. Dá pra ver melhor.

III.A045	Eu acho que um texto em inglês, a pergunta teria que ser inglês, lógico. Mas se tem uma pergunta em português, tá testando muito mais a tradução sua. Entendeu? Não o conhecimento seu com a língua inglesa. Mas sim uma tradução sua.
----------	---

III.A061	O segundo... tipo, é um texto informativo... que tá pedindo pra você falar um... fazer uma redação não fugindo desse tema... falando sobre isso... em português.
P	O que você achou?
III.A062	Eu não sou muito a favor de dar um texto em inglês e pedir em português. Entendeu? Eu acho que deu em inglês, pede em inglês. Deu em português, fala em português. Por isso que eu já nem comentei muito sobre ele.
P	É o tipo de coisa que você não faz no seu dia-a-dia...
III.A063	Não... não faço.
P	... ter que ler em inglês... e escrever em português. Você escreve em inglês mesmo.
III.A064	Ahã.

III.A079	Esse daqui (Tarefa 2)? A tradução. Tradução , com certeza.
P	E mais alguma coisa?

III.A080	Tradução e o seu desempenho em fazer uma redação . Literalmente, que você vai ter que fazer uma redação. Então você tem que ter conhecimento também da língua portuguesa , pra você fazer uma boa redação.
P	Você acha que o seu estilo de escrita interfere?
III.A081	Com certeza. Por isso que eu sou muito mais esse (Tarefa 3). Se dependesse de mim, pra fazer esse daqui (Tarefa 2)... já que já tem que escrever uma redação, eu já ia ficar meio assim... porque eu não sou lá uma das melhores... redatoras... mas pra escrever um e-mail, é uma coisa mais informal... pra um amigo... pra mim, eu acho que já seria muito mais fácil.

III.A084	Com certeza, avalia melhor. Eu tava pensando o seguinte agora... você tava falando dos dois, esse é português (Tarefa 2) e esse é inglês (Tarefa 3)... tá, só que esse (Tarefa 2) é muito mais formal que esse (Tarefa 3) ... entendeu, eu sou muito mais o inglês informal do que o formal .
----------	---

P	E pensando agora nos tipos de tarefas... esse aqui (Tarefa 1), é múltipla escolha, esse (Tarefa 2) é o resumo e esse (Tarefa 3) é o e-mail... você me falou que a 3 você acha que faz mais frequentemente na sua vida. E esse aqui você já fez, o resumo?
III.A086	Já.
P	Em português de um texto em inglês?
III.A087	Na escola, nunca no trabalho, nem na faculdade.
P	Só numa situação de aprendizagem de língua.
III.A088	É, nunca de sobrevivência nem de necessidade.
P	Não de vida real
III.A089	É, não de vida real, só de escola.

III.B032	E esse Segundo (Tarefa 2), sobre a obesidade e o fato de fumar... é um texto mais técnico, tem bastante termos científicos aqui, falando de cromossomos, de telômeros, de encurtamentos... um texto biológico científico... mais restrito, né... seria um nível mais elevado, seria mais um exame... talvez TOEFL, não sei... mas com certeza bem mais... se você fizesse uma comparação... esse texto aqui seria bem mais exigente pra pessoa. Também tem que levar em questão a pessoa que tá lendo, se for uma pessoa da área de saúde, de biologia, vai ler um texto desse mais fácil. Mas pra uma abordagem geral, esse texto aqui é de um nível mais elevado, principalmente essa questão aqui, de fazer o resumo em português... tem várias palavras aqui que não são de uso comum em inglês, então a pessoa talvez tivesse uma dificuldade maior pra traduzir .
----------	---

III.B036	Já esse aqui (Tarefa 2), fazendo uma mera tradução ... eu acho que... eu sempre aprendi inglês não traduzindo, mas pensando em inglês mesmo, entendeu... então eu sou contra traduções de textos. Eu sou a favor de interpretações com perguntas em inglês e respostas em inglês. Mesmo que a sua resposta tenha erros de gramática, mas se você conseguir passar a idéia... é o que vale. A gramática, você pega um livro um de gramática, estuda, faz um curso de um ano e aprende. Eu acho que o que falta aqui seria... poderia fazer um resumo em inglês mesmo...
----------	--

P	E o (Tarefa) 2? Seria adequado como avaliação de leitura?
III.B050	Acho que sim. Porque para você traduzir o texto, você tem que ter entendido bem o texto. Também não adianta traduzir palavra por palavra, porque não vai ficar uma coisa com sentido, então acho que ter entendido o texto é básico pra conseguir fazer esse exercício.
P	Tá. E você faz esse exercício (Tarefa) 2 no seu dia-a-dia? Digo... fazer resumos em português, partindo de um texto em inglês... é uma coisa que você normalmente faz?
III.B051	Sim.... acho que sim... sim, eu faço. Porque tenho vários textos, tenho muita literatura em inglês pra ler na minha área, pra fazer minha dissertação. Só que minha dissertação eu to escrevendo em português, né, então pra referenciar no meu próprio texto, eu tenho que traduzir. Nem sempre eu cito, sabe, coloco ao pé da letra, igualzinho, mas eu tenho que ler, entender e reproduzir aquilo no meu texto, resumindo também. Então é uma coisa que eu faço, sim, bastante.
P	Então você está acostumado a fazer isso, né?.
III.B052	Sim, estou.
P	E você acha que... por estar acostumado a fazer isso... fazer isso numa prova seria mais fácil para você?
III.B053	Então, é aquilo que eu disse... estou acostumado e acho que não seria complicado fazer, mas com esse texto (da Tarefa 2) aqui, acho que seria mais difícil. Porque é aquilo que eu já falei, dos três, esse é o tema mais complexo, tem palavras mais difíceis...
P	Tá, mas imagina que não fosse um texto que você considere muito difícil então. Imagine que fosse com um outro texto, assim, intermediário, sem essas palavras difíceis.
III.B054	Bom... então, sim, seria mais fácil. Principalmente se fosse com um tema da minha área, que eu já conheço os termos, as expressões, já teria mais jeito pra fazer. Mas mesmo que não fosse da minha área, se fosse um tema assim... conhecido, difundido... mais banal mesmo, mais corriqueiro... aí seria fácil fazer o resumo.

P	E o que você acha que o corretor mais iria valorizar durante a correção na (Tarefa) 2?
III.B056	Acho que a tradução... se a pessoa foi capaz de traduzir bem, mantendo o sentido... sem pegar palavra por palavra, mas mantendo o sentido daquele texto.

P	E você já disse aquela hora, é o segundo (Tarefa 2) que tem mais a ver com o que você faz no dia-a-dia?
III.C023	É, o segundo... porque... o que eu tenho feito aqui, né... atualmente. Porque como eu to fazendo uma dissertação... então tudo que eu escrevo, como resumos, como apontamentos, eu faço em português... então todas as minhas dúvidas, como eu to... raciocinando em português... que inclusive é uma das coisas que eu achei importante quando eu morei fora... não é o seu conhecimento específico do inglês, mas a maneira como você pensa. Se você pensa em um e traduz pra outro... não fica tão bom. Então como eu to pensando em português, era como se o meu problema, a minha questão, tivesse sido formulada pelo meu pensamento, que é em português (...).

III.C025	Depende muito da ocasião. Se é uma ocasião que eu preciso fazer um texto em inglês mesmo, é o terceiro. É mais raramente, mas tem que fazer. O mais usual seria mesmo o segundo (Tarefa 2)... porque a maioria dos textos que eu to usando como referência, que eu tenho que ler, são em inglês. Então... mesmo que eu faça qualquer coisa, faça referências pra dissertação, vai ser em português...
----------	--

	então seria o segundo (Tarefa 2) mesmo o mais parecido com o que eu faço .
III.C029	O tipo de escrita da resposta? Sim. Mais no terceiro que no segundo. Pelo meu ver, o resumo que eu fiz em português... não vai ser de interesse se eu escrever um tempo verbal errado em português . Ou sujeito que não tá batendo com o verbo... (...).
P	Embora você considere a Tarefa 3 a mais difícil, você acha que ele se propõe a avaliar leitura... e realmente consegue fazer isso... de um jeito... de um jeito bom... ou você acha que por ser mais difícil ele acaba... saindo um pouco...
III.C032	Acho que não por ser mais difícil... acho que o que poderia atrapalhar um pouco a avaliação seria... a correção estar muito voltada pra COMO eu escrevi o inglês... do que pra O QUE eu escrevi. Porque o segundo fica bem claro que ele vai avaliar O QUE eu escrevi . (...).
III.C033	O segundo. Porque... é basicamente... muito do que a gente faz é usar um texto em inglês como referência... e usar o que eu aprendi disso pra fazer um processamento, por exemplo... e daí fazer uma dissertação sobre isso.

3.3.6. Tarefa 3

Com relação à Tarefa 3, diferentes percepções também foram mostradas, porém de maneira mais uniforme dentro das diferentes faixas de proficiência.

Para os sujeitos I.A, I.B e I.C, de proficiência básica, a Tarefa 3 é distante dos usos de LE que estão habituados a realizar. Embora não tenham achado o tema muito alheio, a tarefa de escrever um e-mail em inglês não é comum, apenas em português. Isso contribui para que a tarefa não seja totalmente estranha para esses sujeitos, que relataram serem capazes de realizá-la, mesmo considerando-a mais difícil. Essa dificuldade se deve à exigência da escrita em inglês, considerada pelos sujeitos como um fator de dificuldade, não apenas por, sendo uma questão aberta, não dar “pistas” de interpretação do texto, mas por exigir a produção numa língua que eles consideram não dominar suficientemente para uma atividade desse tipo.

Assim, como avaliação de leitura, esses três sujeitos demonstraram não considerar a Tarefa 3 adequada, pois, além da alta exigência, a escrita poderia impedir que a capacidade de leitura fosse demonstrada para o examinador. Como já mencionei anteriormente, seria a contaminação da avaliação (Weir 1990) que turvaria os resultados da avaliação.

Os sujeitos II.A e II.B, demonstraram percepções diferentes. Enquanto que para II.A a atividade não se parece com as atividades que desenvolve usando a LE, para II.B ela é muito próxima do seu uso de inglês no dia-a-dia. No entanto, ao contrário do que esse dado poderia levar a supor, II.A considera a Tarefa 3 adequada como avaliação de leitura, devido ao fato de ela, segundo sua percepção, não estar tão calcada ao texto original quanto a Tarefa 2. Ela exigiria elaboração de elementos novos e, justamente por considerá-la mais exigente, o sujeito crê que ela demanda um nível de compreensão do texto mais profundo e, por isso, seria melhor como medida de avaliação. Já II.B, como mencionei anteriormente, percebe o aspecto da contaminação da avaliação na Tarefa 3 e não atribui um caráter positivo a isso. Os dois sujeitos acreditam que a escrita é avaliada integradamente com a leitura nessa tarefa, mas enquanto que II.A percebe isso de maneira positiva, II.B tem reservas quanto a essa forma de avaliação.

Quanto à faixa de proficiência avançada, as percepções sobre a autenticidade variaram um pouco. Os sujeitos III.A e III.B perceberam a tarefa como altamente semelhante ao uso de inglês que realizam no seu dia-a-dia, enquanto que III.C relatou que embora esse tipo de atividade eventualmente esteja presente no seu uso cotidiano de inglês, ela existe em alguns momentos. Não foi a atividade que III.C considerou mais autêntica, mas mesmo assim o sujeito endossou as percepções dos sujeitos da faixa de proficiência avançada, considerando a tarefa, embora exigente, a melhor dentre as três avaliações *para se avaliar proficiência*. Os três sujeitos consideram importante avaliar proficiência globalmente, sem compartimentar as habilidades, e por isso percebem que leitura e escrita são avaliadas na Tarefa. Assim como a percepção demonstrada por II.A, do grupo intermediário, os sujeitos do grupo avançado parecem considerar que o produto da Tarefa 3 deve conter aspectos que somente a reescrita do texto original não contempla, e atribuem a esses aspectos uma maior exigência não apenas com relação à escrita em si mas também ao entendimento necessário do texto para realização da tarefa.

I.A032	Escrever em inglês, eu acho que... é... dificulta bastante , porque na verdade pra escrever em inglês... é, no meu caso, eu demoro bastante, porque... é... bom, se eu for querer escrever rapidamente, eu vou escrever certamente muitas coisas erradas, então teria que voltar pro texto e tentar corrigir meus próprios erros. Acho que isso demanda mais tempo, e acaba dificultando o teste.
--------	---

P	E... qual das tarefas... dentre a 1 a 2 e a 3... você acha que se parece mais com uma situação em que você usaria o inglês na sua vida?
I.A052	É... bom então... eu acho que a terceira tarefa do e-mail pode ser uma atividade bastante corriqueira na vida de muitas pessoas. Na minha não é. Não tenho ninguém com quem eu me comunique em inglês por e-mail. Mas eu acho que pode ser uma atividade corriqueira, mas não é o meu caso. Então considerando o meu dia-a-dia, eu acho que a segunda atividade também é uma atividade que eu realizo mais no meu dia-a-dia, no meu trabalho.

P	E a (Tarefa) 3...
I.A054	A (Tarefa) 3, eu acho que ela pode ser uma atividade rotineira, mas não é o meu caso. Mas eu acho que de várias pessoas que, por exemplo, que trabalham comigo, tem o mesmo tipo de trabalho, usam com uma certa frequência.
P	Escrever e-mails em inglês...
I.A055	É, em inglês.
P	Você escreve e-mails em português pra amigos seus? Desse tipo assim, em português?
I.A056	Sim, em português, sim.

P	Tá. Agora finalmente, o último... o que você achou desse último (Tarefa 3)?
I.A112	Então... parece mais fácil. Mas... é... porque pede pra escrever um e-mail... e aí, na verdade, você... você usa coisas que você leu no texto pra você escrever esse email, né. É... eu acho que parece interessante por ser uma atividade que pode acontecer no dia-a-dia de uma pessoa, mas, assim... escrever em inglês eu acho mais difícil. Então... como já tem... por exemplo, agora... mais de 6 meses que eu não estudo inglês... e como já não tenho muita... é... muito estudo da língua... é... eu sempre fico pensando, ah, o tempo verbal é esse mesmo?... não é?... e aí eu acho que acaba... o texto acaba... o texto fica com mais erros, escrevendo em inglês... é... bom, escrever em inglês é mais difícil, então acho que é esse o ponto.
P	Você acha por exemplo que... esse é uma ferramenta adequada de avaliação de leitura?
I.A113	Então, de leitura... não necessariamente, né, porque aqui você tá avaliando se a pessoa sabe... bom, você vai avaliar se ele leu o texto pelo próprio enunciado da tarefa, pelo o que tá pedindo... é... aí a pessoa tem que usar conceitos que foram ditos no texto... então... é... isso mostra que ela conseguiu ler e entender o texto... mas aí ele tá avaliando escrita também, imagino... porque... às vezes a pessoa pode saber... ler, ler numa boa os textos, consegue interpretar, e na hora de escrever tem uma outra dificuldade né. Que não é o caso justamente essa, quanto mais tempo você fica sem estudar, as dúvidas aparecem mais frequentemente sobre... gramática mesmo... né... então acho que aí é um texto que pode de repente a... pra um corretor, é... se a pessoa que tá fazendo esse teste tem uma dificuldade maior na escrita, pode ficar difícil de passar pro corretor, por exemplo, que ele entendeu o texto...

I.A115	Então, eu acho que ele avalia, sim. Mas... é... eu acho que o fato de ter que escrever em inglês pode dificultar... dependendo do grau de dificuldade que a pessoa tem pra escrever em inglês... eu acho que... pode atrapalhar também a... que o corretor perceba que ele entendeu o que tava escrito no texto em inglês.
--------	---

P	Mas você acha que serve como uma medida de avaliação?
I.A116	Então, eu acho que ele avalia, sim. Mas... é... eu acho que o fato de ter que escrever em inglês pode dificultar... dependendo do grau de dificuldade que a pessoa tem pra escrever em inglês... eu acho que... pode atrapalhar também a... que o corretor perceba que ele entendeu o que tava escrito no texto em inglês.

P	É... o que o corretor ia prestar mais atenção.
I.A118	Ah, tá. Então assim... a solicitação foi essa. Ah... eu acho que... eu acho que ele vai prestar atenção se... bom, primeiro se... é... a pessoa conseguiu entender o enunciado da questão, conseguiu fazer a atividade... se conseguiu colocar o que foi solicitado, por exemplo... é... você vai dizer pra pessoa o que você descobriu desse assunto e dar conselhos práticos pra ela... então acho que ele vai se atentar se... pro e-mail, né... se esses tópicos solicitados no enunciado da questão foram contemplados na escrita, desse e-mail. Acho que ele deve se atentar pra isso... mais do que qualquer outra coisa.
P	Você acha que ele iria , por exemplo... se atentar ao estilo de escrita... isso ia contar alguma coisa na avaliação...?
I.A119	Bom, eu acho que... bom, se é um e-mail... é... ele deve... é... considerar que, bom, se ele tá... penso eu, né... se ele tá pedindo, assim como você escreve uma carta no vestibular, ou uma dissertação, que tem uma estrutura própria... um e-mail também tem uma estrutura informal, na maioria das vezes. Principalmente se é pra um amigo, né. É um texto informal. Então possivelmente... acredito que ele esteja esperando isso também.
P	E você acha que, por exemplo, se você escreve uma carta muito formal, isso seria descontado... seria descontado pontos na avaliação... ou não?
I.A120	Então, tem duas coisas, né. Se é só pra avaliar leitura, eu acho que o estilo do texto não deveria ser considerado. Mas... é... mas... não sei, agora assim tá pedindo um e-mail, pra um amigo, eu não sei se ele tá esperando que eu escreva pra um amigo como EU acho que é... que é um estilo informal... não sei se isso pode contar.
P	Você acha que isso não fica muito claro nas instruções...
I.A121	É... a princípio eu escrevi um e-mail pra... como se eu estivesse escrevendo pra um amigo, estilo totalmente informal. E... ou, acho que em partes, informal... mas eu acho que não fica claro se... se é pra fazer exatamente nesse formato, ou não, mas eu acho que isso também não deve ser... o que tem mais importância nesse tipo de teste.

P	E (que traço importante o corretor iria analisar) no (Tarefa) 3?
I.B038	Acho que é o entendimento do texto em si.
P	O jeito que a pessoa escreve... interfere em algo?
I.B039	Interfere.
P	Como?
I.B040	Tem que usar as palavras corretas... não pode usar palavras muito... populares assim, sabe? Tem que ser a palavra certa. Porque como é um texto assim... você vai ter que usar palavras mais... não mais sérias, também não precisa ser palavras difíceis... mais certas mesmo.

P	O (Tarefa) 3 serve como avaliação de leitura?
I.B041	Sim. Se ela teve que ler, entender, e depois escrever de novo... então... acho que sim. De leitura e escrita.
P	Entre as Tarefas 2 e 3, qual é a mais fácil?
I.B042	A (Tarefa) 2. A (Tarefa) 3, por ter leitura e escrita, é a mais difícil.

P	A (Tarefa) 3 é parecida com seu uso de inglês no dia-a-dia?
I.C047	Não... eu escrevo e-mails, mas não em inglês.
P	Você acha que o fato de não ter essa prática deixaria essa tarefa mais difícil?
I.C048	Sim, com certeza.
P	Das três tarefas, essa (Tarefa 3) seria mais complicada de fazer então?
I.C049	Sim, das três, a mais complicada.
P	E você acha que, pra avaliar leitura, ela (Tarefa 3) é uma... prova, uma... ferramenta... boa? Ela avalia bem a leitura que a pessoa faz em inglês?
I.C050	Eu acho que avalia, porque pra escrever o e-mail, a pessoa tem que ter entendido bem o texto, pelo que ta pedindo aqui no enunciado. Mas não é SÓ a leitura, tem a escrita também. Aí, se você tem uma dificuldade maior pra escrever, como é o meu caso, você sai prejudicado, entendeu? Então tem leitura aí no meio, mas não é só isso. Por isso que eu acho que se é pra avaliar leitura, a melhor é a (Tarefa) 2, que exclui essa parte da escrita.

P	É o tipo de tarefa que você iria melhor.
II.A040	Que eu iria melhor. Que eu teria maior facilidade. E o terceiro é o mais complicado, que eu achei. Que é mais pesado, e tal... chato, né? (risos) Chato, e escrever o e-mail, né, pra uma terceira pessoa... o que já dificulta um pouco.
P	Você acha que o terceiro tem um nível de exigência maior?
II.A041	Maior. Pra mim seria.

II.A046	Qual iria ME avaliar melhor?
P	Avaliar a capacidade de leitura. Você acha que seria um instrumento acurado.
II.A047	Acho que o terceiro...
P	Terceiro? Por que?
II.A048	Porque você teria que... CONTAR... o que você entendeu do texto. Ele fala, manda um e-mail pra uma pessoa que sofre daquele problema, você relatando o artigo que você leu... sugerindo coisas, tal... então, acho que teria uma dificuldade maior, você teria que pensar mais, pra ter um entendimento bom do texto.
P	Você acha que por ser mais difícil... ele seria uma ferramenta mais adequada?
II.A049	É, pois ele exigiria mais.

P	E essa terceira? Qual você acha que é o traço mais importante que ele vai olhar no e-mail?
II.A061	Essa terceira eu acho que é a mais... complicadinha, porque além da leitura entra a escrita , né... envolve, tal... porque a segunda ainda você pode fazer um jogo de montagem ali, eu acho... pode encaixar... mas a 3 exige mais... estaria contando pra outra pessoa...
P	E então... o que você acha que ele... que traço você acha mais importante?
II.A062	Então, aí eu acho que teria duas avaliações, tanto a do entendimento, se ele conseguiu passar aquilo... quanto da escrita , né? Que vai pesar, né... se a escrita não estiver boa, você não vai entender nada...
P	Você falou isso, que aqui na 2 dá pra encaixar... o que você quer dizer com isso? Esse aqui possibilita um uso de estratégias...
II.A063	É, diferente desse 3. Porque um resumo... você vê, aqui [Tarefa 3] você estaria escrevendo pra uma outra pessoa... uma coisa nova , você não... não pode deixar, como que posso dizer...
P	Parafrasear? Você não pode parafrasear?
II.A064	Nesse terceiro? Acho que muda um pouco , né?... A dificuldade é maior.

II.B033	E por último a tarefa 3... seria um artigo pra estar aconselhando a pessoa para estar usando de métodos pra... ta... diminuindo a influência do trabalho sobre ela.
P	Dos três testes, qual reflete as atividade de uso de inglês na sua vida... no seu dia-a-dia?
II.B034	Escrever um e-mail em inglês (Tarefa 3)... exatamente.
P	É o que se aproximaria mais?
II.B035	Sim, é o que se aproxima mais das atividades que eu desenvolvo . Hoje em dia, se você for ver, em todos os lugares o e-mail é amplamente utilizado.
P	Você acha que o (Tarefa) 3 sendo mais parecido, seria mais fácil?
II.B036	Acho que pela questão da prática... acho que é igual aprender a andar de bicicleta. A prática torna mais simples , mas não impede de você... também trabalhar bem em cima das outras coisas. Às vezes você encontraria uma dificuldade maior, mas não creio que seja... de absoluta certeza... que você só vai conseguir trabalhar com aquilo que você já tem prática . Isso depende muito mais da determinação de cada um. Das três é a que exige mais, mas acho que treinando dá pra fazer.

P	E a Tarefa 3... é adequada como... com o avaliação de leitura?
II.B040	Um pouco menos que a (Tarefa) 2, mas ainda assim é boa, porque... pra você conseguir passar pra pessoa, por e-mail, as dicas, o que ela precisa fazer... você tem que ter desenvolvido uma boa leitura, entendido o contexto... caso contrário você não vai conseguir passar o que tá sendo exigido.
P	O fato de estar integrado leitura e escrita, isso influencia a avaliação da leitura?
II.B041	Sim, pelo fato de que as duas formam um conjunto... ela daria uma idéia mais global do conhecimento da pessoa, não fica tão restrito.

P	E (o que o corretor vai valorizar mais) na (Tarefa) 3?
---	--

II.B047	No (Tarefa) 3... o mais importante... acredito que.. na avaliação... acho que seria mais a criatividade e as idéias que você foi capaz de mencionar mediante o problema que você encontrou... Além da parte de desenvolvimento de vocabulário, de gramática... de escrita , né... de vocabulário... é se você foi capaz de sugerir , se as sugestões... acho que as sugestões seriam o que ele ia avaliar na hora de corrigir o texto . É totalmente diferente do (Tarefa) 1, que não precisa fazer uma análise daquilo... daquela resposta que a pessoa deu.
---------	---

III.A050	Primeiro, você tem... o terceiro (Tarefa 3) eu acho que você é mais atividades assim que eu pelo menos usaria. Que é você escrever um e-mail para uma pessoa, dizendo o que você acha, aconselhando, dando a sua opinião... a respeito do problema, que no caso seria... é... um estresse causado por um trabalho, tudo mais... mas podia ser qualquer outro tema... você aconselhar um amigo, ou aconselhar, falar sobre um ponto de vista de trabalho... sobre o que for. No e-mail, falando seu ponto de vista, falando o que você acha.
P	E você acha que esse estilo de... tarefa... pedir para escrever um e-mail... isso é uma coisa que você faria no seu dia-a-dia?
III.A050	Faria. Com certeza. Já fiz muito.
P	Em inglês?
III.A051	Em inglês.

P	E você... então acredita... você disse que acha que essa é a mais parecida com o que você faz no seu dia-a-dia... e eu queria saber se você acha que essa é a melhor tarefa para avaliar.
III.A056	Pra avaliar o conhecimento de inglês de uma pessoa?
P	É.
III.A057	Eu creio que sim. Mas por exemplo, esse primeiro texto aqui, olha (Tarefa 1)... é com certeza o exemplo de uma prova de inglês. Assim que a gente é cobrado. Mas na verdade, na vida real, é só como essa terceira tarefa que a gente é cobrado.
P	Tá, então você acha que o primeiro (Tarefa 1) é mais freqüente... e o terceiro (Tarefa 3) é mais parecido com a vida real.
III.A058	É, isso mesmo. Ahã.
P	Então... entre o freqüente e o que é parecido, qual você acha que é uma ferramenta mais adequada?
III.A059	Pra trabalhar o conhecimento de inglês?
P	Isso.
III.A060	Com certeza, esse daqui, o terceiro (Tarefa 3).

III.A066	Ah, eu não achei que tá... esse segundo, eu achei que tá bem longuinho, mas não tá... aquele texto cansativo... aquele texto que chega no final você já não sabe mais o que tá falando... entendeu? Mas, por exemplo, eu acho que o que tem uma história mesmo, mais vida real, mais realidade, é o terceiro (Tarefa 3) . No entanto que você tá lendo, você sabe do que tá falando... você não vai esquecer, porque tem uma continuidade. Tá falando sobre um assunto.
----------	--

P	E desses três... qual é o tipo de texto que você acha que você lê mais frequentemente?
III.A071	O (Tarefa) 3.
P	O (Tarefa) 3, também.
III.A072	O (Tarefa) 3.

P	Não, não, assim... fazendo o terceiro... o que você acha que você tem que demonstrar pro corretor... pra demonstrar pra ele que você sabe inglês?
III.A076	Que eu sei inglês?
P	É.
III.A077	Não fugir primeiro do tema. Escrever, por exemplo... o terceiro é pra você... pra você falar, pra você dar um conselho pra um amigo. Ele tá sofrendo com isso no trabalho... tá estressado... tudo mais. Pode ficar doente, ou não, tá. É... acho que pra você escrever... você vai ter que aconselhar ele, tudo a respeito desse texto, nunca fugir do texto. Acho que é isso, aí nunca fugir do tema.

P	Você acha que o seu estilo de escrita interfere?
III.A081	Com certeza. Por isso que eu sou muito mais esse (Tarefa 3). Se dependesse de mim, pra fazer esse daqui (Tarefa 2)... já que já tem que escrever uma redação, eu já ia ficar meio assim... porque eu não sou lá uma das melhores... redatoras... mas pra escrever um e-mail, é uma coisa mais informal... pra um amigo... pra mim, eu acho que já seria muito mais fácil.

P	Entendi. Tá, e você acha que o fato de, por exemplo, essa aqui... a número 3 e a número 2, que essa aqui é a em português e essa é a inglês... mas você acha que o fato de ter leitura e escrita junto... isso dificulta?
III.A083	Com certeza. Dificulta , pq além de você ter o reading, você tem que ter o listening, tem que ter o writing... primeiro, você vai ter que LER... ENTENDER... você vai ter que ouvir você lendo, entendendo... pra depois você escrever. Acho que isso dificulta muito mais, se fosse só ler...
P	Mas mesmo sendo mais difícil você acha que é mais adequado?
III.A084	Com certeza, avalia melhor. Eu tava pensando o seguinte agora... você tava falando dos dois, esse é português (Tarefa 2) e esse é inglês (Tarefa 3)... tá, só que esse (Tarefa 2) é muito mais formal que esse (Tarefa 3)... entendeu, eu sou muito mais o inglês informal do que o formal.

P	Você acha que o fato de você fazer atividades parecidas com essa número 3 influencia quando você tem que fazer uma prova do tipo dessa? O fato de você ter que fazer isso constantemente facilita?
III.A092	Com certeza vai facilitar , porque já estou acostumada a fazer um e-mail aconselhando, tudo mais... me facilitaria muito mais fazer essa que fazer um teste... que a princípio seria mais simples, exigira menos do meu inglês do que o e-mail... e pra mim seria muito mais difícil fazer um teste desse, com múltipla escolha.

P	Então qual dessas tarefas... fazer múltipla escolha (Tarefa 1), fazer um resumo (Tarefa 2), ou
---	--

	escrever um e-mail (Tarefa 3)... é mais parecido com o que você faz com o inglês no seu dia-a-dia?
III.B030	O e-mail (Tarefa 3), com certeza. Porque aí tem que ler um... um texto, que eu acho que é de internet... eu leio muito na internet... mandar um e-mail... que eu também mando e-mails pros meus amigos que não são daqui, então tem que ser em inglês... o resumo eu faço bastante também, mas acho que menos, e mais por obrigação também. Já a múltipla escolha eu faço quando tenho alguma prova, tenho que estudar, me preparar... aí pego provas, exercícios, e me preparo fazendo antes.

III.B033	Esse terceiro aqui (Tarefa 3), basicamente o que eu poderia dizer... é que é o mais difícil , eu acho... que eu achei. Não o mais difícil, seria o mais complicado. Você dar conselhos... pra como evitar que seu trabalho prejudique a sua saúde. Se você for usar os argumentos que estão usando aqui, você vai estar necessariamente fazendo uma tradução. Então... isso aqui é meio complicado de falar um pouco... talvez se eu lesse ele com mais calma, talvez eu pudesse escrever melhor. Mas nessa ordem que eu citei, esse terceiro texto aqui (Tarefa 3) foi o mais... mais atípico de todos, mas de menos dificuldade do que esse aqui... o 2 (Tarefa 2)... esse (Tarefa 3) eu achei um intermediário. Interessante o assunto também... só que meio técnico, voltado assim pra parte de biológicas... tem umas palavras aqui, <i>epinephrine</i> , acredito que seja um hormônio, uma proteína, ou uma... enzima... isso aqui é um texto biológico. Dependendo do público que vai ter acesso a um texto desses pra ter conselhos, pra dar conselhos... acho que seria meio restrito, num tipo de exame de vestibular... se fosse pra um exame, por exemplo, pra área de saúde... pra um médico... acho que talvez os exames devessem ser mais voltados pra cada área de atuação, porque cada profissional vai ler uma coisa mais voltada pra sua área. (...)
P	Então você achou o tema desse (Tarefa 3) difícil.
III.B034	Difícil não... achei meio restrito. Assim... menos abrangente. Difícil não. Assim... dá pra fazer. Ler e entender os três, eu entendi. Só que assim... o que eu teria mais dificuldade pra fazer seria esse (Tarefa 3). Eu achei assim o mais, entre aspas... “chatinho”, cansativo de fazer.
P	Ahã. E com relação aos exercícios que vieram junto com eles? Qual a sua impressão geral, por exemplo, sobre esse 3 (Tarefa 3)?
III.B035	Esse 3 aqui... seria pra escrever em português ou em inglês?
P	Inglês.
III.B036	Inglês, tudo bem. Seria uma dificuldade a mais. Eu acho que o certo é isso mesmo. Já que o texto é um texto que não é tão fácil... a pessoa, o aluno em questão lendo e escrevendo em inglês uma coisa coerente... o resultado você vai ter a prova que ele realmente compreendeu e conseguiu se expressar (...).

P	E qual dos três é mais parecido com o que você faz no dia a dia?
III.B043	A (Tarefa) 3.
P	E isso tornaria mais fácil?
III.B044	Não, acho que a (Tarefa) 1 é a mais fácil
P	Então, mesmo estando acostumado, a (Tarefa) 3 seria a mais difícil?
III.B045	É, porque exige mais de você. Exige que leia ,entenda, depois reescreva sobre o tema... que não é um tema assim, comum... já a 1 é um tema mais facinho de fazer.

P	Mas tirando a parte do tema... suponha que a Tarefa 1, de alternativas, fosse com o texto da (Tarefa)3...
III.B046	Hã... o que tem?
P	Mesmo assim a (Tarefa) 1 seria a mais fácil então? Porque você disse que o problema é o texto... mas vamos pensar na parte da tarefa mesmo. A tarefa de escrever um e-mail e a tarefa de múltipla escolha.
III.B047	Ah, tá... então, acho que ficaria mais difícil, mas ainda não tanto quanto o e-mail . Se a Tarefa 1 fosse com um texto mais difícil, ficaria mais complicado, lógico... e a mesma coisa a (Tarefa) 3, se fosse com um texto mais fácil, ficaria mais fácil... mas mesmo assim, você ter que ler, entender e tudo... e depois escolher uma alternativa ainda é menos complexo do que ter que pensar no e-mail, em como vai passar as sugestões, em que palavra vai usar... na (Tarefa) 1 não tem que pensar em nada disso, então mesmo com um texto mais complexo, ainda assim acho que não seria um exercício tão difícil quanto a (Tarefa) 3 .
P	Vc acha que o (Tarefa) 3 seria um teste adequado de leitura?
III.B048	Não só de leitura, mas de leitura e escrita, né? Claro que tem leitura ali, mas não é só isso , acho que é as duas coisas... o que, na verdade, acho que é o melhor , fica uma coisa mais global.
P	Mesmo sendo mais difícil?
III.B049	Sim.

III.B055	Nessa tarefa 3 aqui, que é solicitado que o aluno dê conselhos para um amigo... eu, se fosse o corretor, iria levar em consideração o que ele propôs... as idéias dele... a posição dele, tendo o cuidado de perceber que não foi uma coisa COPIADA do texto, mas que foi uma coisa que partiu da idéia dele , que ele foi capaz de gerar períodos, de gerar frases, ter vocabulário, conseguir ordenar isso numa frase e passar uma idéia. Isso que seria avaliado por mim, não os conselhos em si.
----------	---

P	Comparando Tarefa) 2 e (Tarefa) 3, qual é melhor pra avaliar leitura?
III.B057	Depende, se for só leitura que o cara quer, é a (Tarefa) 2... mas eu acho que a (Tarefa) 3 é uma atividade mais completa, é a mais ideal . Eu acho. Ela é mais exigente, demanda mais coisas da pessoa, não é só pegar o texto e pronto... tem que pegar o texto e fazer uma coisa a mais, pra isso vai exigir mais do inglês dele... e ele vai poder mostrar mais o seu conhecimento . Por isso pra avaliação, eu acho que a 3 é a mais adequada, porque tem condições de mostrar mais... mais variedade, mais coisas que você pode fazer com seu conhecimento da língua, não fica preso só naquele resumo... que, aliás, é em português, né... quero dizer, se ainda fosse reescrever com suas palavras em inglês, ainda vá lá, mas nem em inglês é, é em português. Já no terceiro, não é que você vai reescrever em inglês, é mais que isso... é um texto diferente, um e-mail, que é uma coisa nova, diferente do texto dado na prova... então exige mais, é melhor .

III.C020	Eu achei o primeiro mais fácil. Mas a sua pergunta é... qual dos três seria mais adequado pra uma avaliação, né?
P	É.
III.C021	Eu acho que seria o terceiro (Tarefa 3).
P	Por quê?

III.C022	Porque quando ele vai desenvolvendo o texto... por exemplo, grande parte dessas perguntas que tem no primeiro (Tarefa 1)... tipo, o que significa essa palavra no texto... já tem embutido no entendimento do texto. Nos três... nos três testes. Só que o terceiro, em relação ao primeiro... é aquela questão, quando eu vou escrever, eu não tenho nenhuma dica de qual possa ser o jeito mais certo . Seriam as múltiplas escolhas. E em relação ao segundo, eu achei o e-mail mais interessante pelo fato da questão ser em inglês . Se eu faço uma pergunta em inglês, e vou responder... e não entendi muito bem a pergunta, todo processo dali pra baixo vai embora. Enquanto que o segundo, como a pergunta é em português... essa etapa do conhecimento, do entendimento da questão... ela fica mais fácil.
III.C029	O tipo de escrita da resposta? Sim. Mais no terceiro que no segundo. Pelo meu ver, o resumo que eu fiz em português... não vai ser de interesse se eu escrever um tempo verbal errado em português. Ou sujeito que não tá batendo com o verbo... então... isso o terceiro eu acho que já levaria em conta . Porque... já que você está escrevendo em inglês... tudo bem que você entendeu o outro, mas pra você conseguir reproduzir isso em inglês é um trabalho a mais, e por isso eu acho ele mais interessante .
P	Você acha mais difícil?
III.C030	Sim. Escrever em inglês, sim.
P	Se você fosse enumerar por ordem de dificuldade, do mais difícil pro mais fácil
III.C031	É (Tarefa) 3, (Tarefa) 2, (Tarefa) 1.. Então, pra mim... o terceiro é mais complexo. Ele envolve um conhecimento de texto... de interpretação de texto... de estrutura, que eu vou escrever, de inglês mesmo . O segundo, é mais de interpretação... e o primeiro, coisas mais isoladas no texto.
P	Embora você considere a Tarefa 3 a mais difícil, você acha que ele se propõe a avaliar leitura... e realmente consegue fazer isso... de um jeito... de um jeito bom... ou você acha que por ser mais difícil ele acaba... saindo um pouco...
III.C032	Acho que não por ser mais difícil... acho que o que poderia atrapalhar um pouco a avaliação seria... a correção estar muito voltada pra COMO eu escrevi o inglês... do que pra O QUE eu escrevi . Porque o segundo fica bem claro que ele vai avaliar O QUE eu escrevi. O terceiro, aí vai depender de quem vai avaliar... ele pode correr o risco de avaliar COMO eu escrevi . Nesse ponto de vista... a interpretação de texto pode ficar um pouco aquém do meu conhecimento de escrita... de estruturas.

Apresentados os dados e a análise efetuada, a seguir, no Capítulo IV, estão as conclusões tecidas a partir dessa discussão, apontando implicações e limitações deste estudo.

CAPÍTULO IV

CONCLUSÕES

Neste capítulo final, retomo a discussão sobre os dados feita no Capítulo III, assim como aponto implicações e limitações do estudo, sob a luz do referencial teórico adotado e com vistas a responder à questão de pesquisa colocada, a saber:

Como se configura o conceito de autenticidade na avaliação de leitura em LE para os examinandos?

Para responder a essa questão, é preciso retomar alguns caminhos apontados pelo arcabouço teórico usado para sustentar o trabalho. Um desses caminhos apontava a relatividade do conceito de autenticidade, que poderia se configurar de diferentes maneiras para diferentes sujeitos.

Sobre esse tópico, um dos pontos mais importantes a ser ressaltado é o da relatividade das percepções acerca da integração de habilidades como um elemento que constitui aparente autenticidade das avaliações para os examinandos. Pode-se notar isso quando comparando alguns dos dados relativos a cada uma das tarefas: alguns sujeitos demonstram ter impressões muito díspares em relação aos outros. Alguns consideram a tarefa válida por ser mais exigente, enquanto que outros a consideram menos válida justamente por isso. O mesmo pode ser dito com relação à percepção do que, aparentemente, eles acreditam que os examinadores pretendem com aquele item: embora pareça haver consenso de que na Tarefa 3, por exemplo, leitura e escrita estão sendo avaliadas conjuntamente, para alguns sujeitos isso não parece “justo”; para outros, esse é o procedimento correto para fins de averiguação da proficiência, mesmo que mais exigente.

Nessas diferentes percepções, um ponto que a princípio não seria levado como fator de análise, o *tema dos textos*, apareceu como uma variável que parece orientar as percepções dos examinandos sobre os itens avaliativos e, conseqüentemente, sobre sua autenticidade, principalmente quando levamos em consideração que alguns sujeitos associaram muito a noção de facilidade percebida nos itens com a autenticidade deles. A

presença de um tema familiar na tarefa parece fazer com que os examinandos a considerem mais autêntica, a despeito do seu formato não se conformar a uma tarefa de uso de língua semelhante ao que realizam habitualmente. Em outros casos, mesmo não sendo um tema familiar, mas um tema que pareça mais agradável aumenta a motivação dos examinandos para a realização da tarefa. No entanto, nem todos os sujeitos expressaram percepções acerca dos temas das tarefas. Por isso, seria necessário um olhar atento para essa questão para maior elucidação dos fatores que desencadeiam essas diferentes percepções e que motivam os sujeitos de diferentes maneiras, o que não foi o objetivo desta pesquisa. Vale a pena ressaltar, entretanto, que esse ponto merece consideração em pesquisas futuras, visto que em alguns momentos, para alguns sujeitos, a variável “assunto do texto” parece mesmo mais importante que o formato da tarefa proposta, atuando na percepção que eles constroem sobre a autenticidade e a validade da tarefa.

Da mesma maneira, um outro sentido atribuído ao conceito de autenticidade pelos sujeitos, não tomado a princípio como um dos sentidos que busquei investigar, parece ser muito relevante na percepção que constroem sobre o item avaliativo. No Capítulo II, discutindo sobre o conceito de autenticidade, procurei associá-la à relação entre tarefas de uso real de língua e tarefas avaliativas, destacando que seu cerne é o quanto que estas últimas (tarefas avaliativas) aproximam-se das primeiras (tarefas de uso real de língua). A Figura 4 abaixo procurava delinear essa idéia:

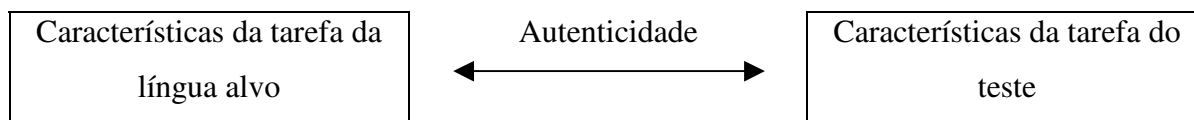


Figura 4: Autenticidade (Bachman & Palmer, 1996, p. 23)

Um outro sentido de autenticidade, no entanto, coexiste com a noção demonstrada na figura nas percepções dos sujeitos: o sentido de autenticidade da tarefa não pelo quanto ela se parece com uma tarefa de uso real de língua, mas pelo quanto ela se parece com uma *tarefa avaliativa*. Ou seja, consoante com o que Alderson (2000, p.53) afirma sobre as avaliações de língua constituírem, por si mesmas, eventos tão autênticos de uso de LE quanto qualquer outro encontrado na vida real dos falantes de uma língua, um

dos sentidos de autenticidade que os sujeitos procuram nos itens não é se eles se parecem com seu uso real de língua, mas sim se elas se parecem com o uso real de língua *feito em situações de avaliação*. Assim, nenhum dos sujeitos atribuiu autenticidade à Tarefa 1 (múltipla escolha) no sentido de se parecer com uma tarefa que fazem habitualmente em suas vidas, mas assinalaram a autenticidade dela como um evento letrado próprio de situações formais de ensino e avaliação. É muito importante assinalar que esse é o tipo de fator que contribui na construção da validade de face do item, e que, portanto, atua na aceitação e crédito que os examinandos atribuem a um determinado item avaliativo – um fator grandemente marcado pelas crenças sobre aprendizagem e avaliação dos examinandos construídas ao longo de sua vivência letrada.

Pode-se dizer que existe uma certa convergência no que se refere às percepções muito relacionadas ao nível de proficiência demonstrado. Sujeitos de proficiência mais elementar parecem ter uma clara preferência por itens que considerem mais fáceis. Neste estudo, os três sujeitos consideraram a Tarefa 2 (resumo) mais parecida com os usos de LE que realizam em suas vidas cotidianas. Isso, no entanto, parece não ser o que governa suas percepções sobre a validade do item. Apenas um dos três sujeitos (sujeito **I.A**) considerou a Tarefa 2 como um melhor instrumento de avaliação de sua proficiência em leitura, exatamente por lhe parecer mais autêntica; a despeito de terem a mesma opinião sobre a autenticidade da tarefa, os outros dois sujeitos (sujeitos **I.B** e **I.C**) julgaram a Tarefa 1 (múltipla escolha) mais adequada como instrumento avaliativo da leitura. **I.A** considerou a Tarefa 1 mais difícil por causa das alternativas, que, segundo sua percepção, poderiam confundir a escolha da resposta correta, julgando, pois, a Tarefa 2 mais fácil; **I.B** e **I.C** consideraram a Tarefa 2 mais difícil, devido ao componente escrito solicitado em sua realização, considerando mais fácil a Tarefa 1. Um ponto, no entanto, é comum aos três sujeitos dessa faixa de proficiência: todos atribuíram maior validade à tarefa que julgaram mais fácil, ou melhor dizendo, com um nível de exigência mais adequado ao nível de proficiência que acreditam ter.

Já com os sujeitos de proficiência mais alta, algo diverso se dá. Dos três sujeitos de proficiência mais avançada, dois (sujeitos **III.A** e **III.B**) julgaram a Tarefa 3 (e-mail) a mais parecida com os usos de LE que realizam cotidianamente, sendo que **III.A** a

considerou mais fácil e **III.B**, mais difícil; um sujeito (**III.C**) considerou a Tarefa 2 (resumo) a mais autêntica e mais fácil, e a Tarefa 3, a mais difícil. Apesar dessas diferenças com relação à autenticidade e facilidade que cada um deles atribuiu, os três sujeitos atribuíram maior validade à Tarefa 3.

Um primeiro olhar para esses dados poderia nos levar à conclusão de que, ao contrário dos sujeitos da faixa de proficiência mais básica, os da mais avançada consideraram os itens mais difíceis mais válidos (já que dois de três sujeitos julgaram a Tarefa 3 o item mais difícil e mais válido), ou que eles consideraram os itens mais autênticos mais válidos (já que, da mesma forma, dois de três sujeitos julgaram a Tarefa 3 mais autêntica e mais válida). A atribuição dessa validade, no entanto, pode estar ligada não apenas à facilidade/dificuldade, ou à autenticidade, percebida nos itens, mas ao que os sujeitos *acreditam ser importante ser contemplado numa avaliação em LE*. Mesmo quando questionados sobre se o item lhes parecia mais adequado para a avaliação da leitura, os três demonstraram considerar a leitura e produção em LE como a forma mais válida de avaliação de leitura, por acreditarem que a LE deve ser avaliada integralmente, e não compartimentada em suas macro-habilidades. Embora apenas um dos sujeitos tenha considerado a Tarefa 2 mais difícil, as falas de todos demonstraram a percepção de que a realização da Tarefa 3 exige interpretação e resignificação dos elementos do texto, ou seja, a tarefa seria mais exigente – para um dos sujeitos, essa exigência não necessariamente tornaria a tarefa mais difícil. Nesse caso, poderíamos dizer que, para esses sujeitos, o conceito de validade estaria ora imbricado com o de “exigência” do item, ora com o de autenticidade.

Essa mesma crença de que a integração de habilidades é o que torna o item mais válido parece orientar as percepções de um dos sujeitos da faixa de proficiência intermediária (sujeito **II.A**), que atribuiu maior validade à Tarefa 3, mas maior autenticidade à Tarefa 2. Nota-se aqui, assim como na análise dos dados do sujeito **III.C**, que a maior integração de habilidades, embora historicamente aponte para o sentido de busca de maior autenticidade na avaliação (ver no Capítulo II a discussão sobre integração de habilidades e sua relação com autenticidade), não é sinônimo de maior autenticidade, e nem de maior validade. Esses sujeitos, consideraram a Tarefa 2 mais autêntica que a Tarefa 3, embora

menos integrada. E justamente a maior integração de habilidades da Tarefa 3 seria reflexo de sua maior validade – e não sua maior autenticidade como índice de maior validade, como poderia se supor.

Abordei essa questão na página 50, quando comentando a questão dos testes diretos e indiretos discutidos por Low (1985). O autor questiona a idéia de maior validade *a priori* das avaliações autênticas e usa como argumento o fato de que tantos e tão variados são os usos autênticos de língua no dia-a-dia que se torna difícil definir balizas de validação para esse tipo de teste. Embora aqui estejamos falando de validade de face – ou seja, não de validade definida em termos rigorosos, como propunha Low (id.) em seu estudo – é interessante notar que o mesmo questionamento se coloca: a atribuição de validade não é feita devido à autenticidade do item, mas sim a outras crenças que os sujeitos tenham, seja sobre a importância da integração das habilidades (caso dos sujeitos de proficiência avançada), seja pela crença que os sujeitos têm sobre a própria proficiência e qual nível de exigência adequado a ela (caso dos sujeitos de proficiência básica).

Pelos dados analisados, a autenticidade parece ser relativa não apenas no que se refere aos diferentes usos de LE, mas também com relação ao nível de proficiência e às visões e crenças sobre língua e avaliação que cada um dos sujeitos tem. Esse caminho de interpretação dos dados é consoante com o que sugere Bartholomeu (2002), quando investigando as percepções dos examinandos sobre vestibulares de inglês e sua relação com o efeito retroativo exercido por eles. A autora indica em seu estudo que essas percepções não são uniformes devido a diversos fatores, entre os mais importantes os valores e crenças com relação a avaliações de LE trazidas das histórias de aprendizagem dos sujeitos.

No mesmo sentido, vemos configurar-se aqui um quadro que mostra que não é a proximidade das tarefas avaliativas às tarefas de uso real da língua – ou seja, a autenticidade da avaliação – que determina a validade de face desses itens, e sim outras crenças, como a semelhança do item à idéia de avaliação que os examinandos têm formada (como já discuti, um outro sentido de autenticidade da avaliação), a importância da integração de habilidades para a constituição de uma avaliação mais “completa”, ou a facilidade e adequação a um determinado nível de proficiência atribuída ao item.

A partir disso, retomamos dois pontos importantes discutidos no Capítulo II, de dois estudos anteriores que apontaram reflexões que vão ao encontro das aqui apresentadas. Um deles é o questionamento de Cumming e Maxwell (1999) sobre a tomada da autenticidade na avaliação, por muitos, como “obviamente uma coisa boa”. As autoras apontam vários elementos que questionam essa premissa, todos versando sobre interpretações equivocadas do que significa autenticidade que podem levar a práticas de validade e confiabilidade duvidosas. A esses fatores, poderíamos adicionar outro: a autenticidade é uma “coisa obviamente boa” também segundo a visão dos examinandos?

Os dados aqui analisados apontam para a direção do “não”. Esse resultado é consoante com o do estudo de Lewkovicz (2000), cujos sujeitos não perceberam a autenticidade como algo importante na avaliação, que permitiria avaliar suas capacidades lingüísticas de uma melhor maneira. Uma limitação apontada nesse estudo, no entanto, era a falta da definição do que era autenticidade para esses sujeitos. Na presente pesquisa, investiguei a importância atribuída à autenticidade pelos sujeitos, mas procurando delinear um conceito de autenticidade a partir de suas percepções. A análise desses dados demonstra alguns ganhos para a compreensão de como se configura esse conceito de autenticidade para esses examinandos.

Uma das questões que assinali anteriormente que torna a investigação das perspectivas dos examinandos importante é a falta de correspondência entre as perspectivas dos diferentes envolvidos no processo avaliativo, os *protagonistas* do processo, sobre uma determinada avaliação. Como já havia afirmado, isso causa problemas tais como o almejar diferentes objetivos dos diferentes protagonistas, prejudicando, assim, o resultado e as influenciando nas conseqüências deles. Um exemplo que podemos tirar dessa própria pesquisa é a percepção que muitos sujeitos demonstraram da Tarefa 2 como uma tarefa de *tradução*, e não de resumo. Tradução e resumo são atividades diferentes, e objetivo da Tarefa 2 era o segundo. No entanto, talvez pela crença de que tradução e resumo de um texto em LE sejam a mesma tarefa, ou que tradução é um reflexo de compreensão de texto, esses sujeitos buscariam realizar a Tarefa 2 com outro propósito, que não o almejado na formulação do item.

Uma das implicações de compreender esse tipo de diferença nas percepções é, em primeiro lugar, a atenção ao problema que isso obviamente configura num processo avaliativo. E esse problema nos leva a pensar não apenas sobre o tipo de práticas letradas esses sujeitos vivenciam em LE, mas mesmo em sua própria língua, e também se eles as relacionam a uma situação avaliativa. A percepção de uma tarefa de resumo como tarefa de tradução coloca em cheque o conhecimento que esses sujeitos têm de diferentes gêneros textuais formais, e o quanto a escola, como instituição responsável pelo ensino/aprendizagem de gêneros escritos, foi eficiente nessa atividade. Este dado nos atenta para o fato de que mesmo examinandos com um alto nível de escolaridade e letramento podem construir percepções equivocadas sobre nossos instrumentos avaliativos, e que é tarefa da escola e dos responsáveis pela formulação e correção de exames externos promover orientação sobre essas possíveis percepções equivocadas.

Dentro da escola, por exemplo, investigar as percepções dos alunos sobre os instrumentos avaliativos pode alertar para elementos causadores de baixo rendimento que poderiam não ter sido considerados importantes – esclarecer o que se espera, e como se espera, de um aluno em uma avaliação é um primeiro passo para evitar percepções equivocadas. Além disso, como já salientaram Cumming e Maxwell (1999), uma maior compreensão do processo avaliativo por parte dos alunos pode levá-los a desenvolver uma atitude mais positiva frente a essa avaliação, trazendo ganhos para o processo de ensino/aprendizagem e da avaliação.

Num exame externo, como já salientei anteriormente, a investigação da validade de face atua no campo de aceitação e adoção do exame como ferramenta avaliativa por examinandos e instituições que dependem desses resultados para tomadas de decisões, como empresas privadas, universidades, agências fomentadoras de pesquisa, órgãos e conselhos reguladores etc. Investigar que percepção os examinandos apresentam sobre o exame pode fornecer pistas sobre o efeito retroativo que esse exame exerce, e se ele é o que esperam seus formuladores. O exame CELPE-Bras, por exemplo, é um exame cujo formato pretende provocar o efeito retroativo de estimular a aprendizagem contextualizada da língua portuguesa, o trabalho com gêneros, a compreensão de textos que vá além da localização e colagem de informações etc. Investigar as percepções que os candidatos têm

sobre o exame leva à reflexão sobre se esses objetivos estão sendo atingidos e se o efeito retroativo do exame tem sido aquele almejado por seus formuladores.

Em ambos os contextos – de avaliação escolar ou avaliação externa – os resultados dessa pesquisa levam a pensar, em primeiro lugar, que, embora a visão de professores e examinadores já tenha mudado bastante com relação a abordagens de ensino, é necessário investigar o quanto as crenças dos alunos encontram consociação nas práticas avaliativas propostas.

Um segundo ponto a ser levantado é o modo como os itens que buscam ser autênticos são apresentados aos aprendizes e examinandos de proficiência mais elementar. Como destacamos quando discutindo as colocações de Widdowson (1998), não são todas as situações “autênticas” de uso da língua que se prestam à aprendizagem de LE; no mesmo sentido, talvez não sejam todas as tarefas autênticas que se prestem à avaliação de todos os níveis de proficiência. Numa situação escolar, o professor deve ter critérios e objetivos bem definidos para sua avaliação, sabendo escolher um tipo de tarefa significativa, que busque integração e autenticidade, adequada ao nível de proficiência de seus alunos, cujos propósitos os alunos não percebam como muito aquém de sua capacidade. A mesma preocupação deve estar presente na formulação de exames externos que apresentam diferentes provas para diferentes níveis de proficiência (como, por exemplo, os KET, PET, FCE etc, de responsabilidade do UCLES). Já no caso dos exames externos que apresentam uma única avaliação e, a partir do desempenho demonstrado, classificam os candidatos em faixas de proficiência (como o IELTS e CELPE-Bras, por exemplo), os objetivos e critérios de avaliação devem estar muito claros, e a grade de correção deve ser muito cuidadosa, buscando prever e abarcar as diferentes ocorrências demonstradas pelos candidatos no exame e saber classificá-las adequadamente para que o efeito retroativo desse exame seja o esperado. No mesmo sentido, para que diferentes percepções dos examinandos possam ser orientadas para o sentido almejado pelos examinadores, não necessitando de reformulações do exame, se não for esse o caso, mas estimulando mudanças nas práticas de ensino e aprendizagem da LE de modo a contemplar as capacidades exigidas pelo exame.

As questões aqui discutidas indicam alguns pontos que já podemos compreender melhor sobre autenticidade, integração de habilidades e percepção dos

examinandos sobre as avaliações de línguas. No entanto, o escopo da pesquisa é limitado, e muitas questões continuam por ser investigadas. É necessário ainda refletir sobre que outros fatores podem estar imbricados nos perfis desses sujeitos, como diferentes histórias de aprendizagem de LE, motivações, experiências na realização de testes e outros, para conseguir delinear melhor o que atua na construção dessa percepção e como ela influencia no desempenho desses examinandos. Esclarecimentos desse tipo nos levarão não apenas à prática mais consciente de nossos papéis de professores e examinadores, mas principalmente à aplicação mais eficiente e proveitosa da avaliação em nossa sociedade.

REFERÊNCIAS BIBLIOGRÁFICAS

- Alderson, J.C. (1979) The cloze procedure and proficiency in English as a foreign language. *TESOL Quarterly* 13 (2), pp. 219-227.
- _____ (2000) *Assessing Reading*. Cambridge: Cambridge University Press.
- Bachman, L.F. (1990) *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press.
- Bachman, L.F. & Palmer, A.S. (1996) *Language Testing in Practice*. Oxford: Oxford University Press.
- Banerjee, J. & Luoma, S. (1997) Qualitative approaches to test validation. In: Clapham, C. & Corson, D. (orgs.) *Encyclopedia of Language and Education, Volume 7: Language Testing and Assessment*, pp.275-287.
- Bartholomeu, M. A. N. (2002) *Prova de língua estrangeira (Inglês) dos vestibulares e sua influência nas percepções, atitudes e motivações de alunos do terceiro ano do ensino médio*. Dissertação de mestrado, Programa de Pós-graduação em Linguística Aplicada, IEL, Unicamp.
- Bradshaw, J. (1990) Test-takers' reactions to a placement test. *Language Testing* 7(1), pp. 13-30.
- Cattell, R.B. (1964) Validity and reliability: a proposed more basic set of concepts. *Journal of Education Psychology* 55, pp.1-22.
- Charge, N. & Taylor, L.B. (1997) Recent developments in IELTS. *ELT Journal*, 51 (4), pp. 374-380.
- Clark, H.H. (2000) O uso da linguagem. *Cadernos de tradução* 9, pp. 49-71.
- Clark, J.L.D. (1975) Theoretical and technical considerations in oral proficiency testing. In: Jones, R.L. & Spolsky, B.(orgs.) *Testing language proficiency*. Arlington, Virginia: Center of Applied Linguistics.
- Cohen, A.D. (1989) Metodologia de pesquisa em Linguística Aplicada: Mudanças e Perspectivas. *Trabalhos em Linguística Aplicada* 13, pp. 1-13.
- Cumming, J.J. & Maxwell, G.S. (1999) Contextualising authentic assessment. In: *Assessment in education* vol. 6 (2) Abingdon, pp. 177-197.
- Hill, C. & K. Parry (1992). The Test at the gate: models of literacy in reading assessment. *TESOL Quarterly* 26 (3), pp. 433-461.
- Hughes, A. (1990) *Testing for language teachers*. Cambridge: Cambridge University Press.
- Ingram, D. (1984) Basic concepts in testing. In: Allen & Davies (1977) pp. 11-37.
- Kaplan, R.B. (1988) Process vs product: problem or strawman. *Lenguas Modernas* 15, pp. 35-44.

- Larsen-Freeman, D. & Long, M.. (1991) Second Language acquisition research methodology. In: *An introduction to second language acquisition research*. London: Longman. 10-51.
- Lee, J.F. (1986) Background Knowledge & L2 Reading. *The Modern Language Journal* 70 (4), pp. 350-354.
- Lewkowicz, J. A. (1997) The integrated testing of a second language. In: Clapham, C. & Corson, D. (orgs.) *Encyclopedia of Language and Education, Volume 7: Language Testing and Assessment*, pp. 121-120.
- _____ (2000) Authenticity in language testing: some outstanding questions. *Language Testing* 17 (1), pp. 43-64.
- Low, G. (1985) Validity and the Problem of Direct Language Proficiency Tests. in: Alderson, J.C. (ed.) *Lancaster Practical Papers in English Language Education, Volume 6: Evaluation*. Oxford: Pergamon Press.
- Lüdke, M. e André, M.E.D.A. (1986) *Pesquisa em Educação: Abordagens Qualitativas*. São Paulo: Editora Pedagógica e Universitária.
- Lumley, T. & Brown, A. (2004) Test-taker and Rater Perspectives on Integrated Reading and Writing Tasks in the Next Generation TOEFL. *Language Testing Update* 35, pp. 75-79.
- Moita Lopes, L.P. (1994) Pesquisa interpretativista em Lingüística Aplicada: A Linguagem como condição e solução. *D.E.L.T.A.* 10 (2), pp. 329-338.
- Moni, K.B., van Kraayenoord, C. & Baker, C.D. (2002) Students' perceptions of Literacy Assessment. *Assessment in Education* 9 (3), pp. 319-339.
- Nevo, B. (1985) Face Validity Revisited. *Journal of Educational Measurement* 22 (4), pp. 287-293.
- Oller Jr., J. W. (1979) *Language Tests at School*. London: Longman.
- Scaramucci, M.V.R. (1990) O Resumo e a Avaliação da Compreensão em Leitura em Língua Estrangeira. *Trabalhos em Lingüística Aplicada* 15, pp. 65-86.
- _____ (1995) *O papel do léxico na compreensão em leitura em língua estrangeira: foco no produto e no processo*. Tese de Doutorado. Campinas: Universidade Estadual de Campinas.
- _____ (2000) Proficiência em LE: Considerações Terminológicas e Conceituais. *Trabalhos em Lingüística Aplicada* 36, pp. 11-22.
- _____ (2004a) Efeito retroativo da avaliação no ensino/aprendizagem de línguas: o Estado da Arte. *Trabalhos em Lingüística Aplicada* 43 (2), pp. 203-226.
- _____ (2004b) The Brazilian Certificate of Proficiency in Portuguese as a Foreign Language – CELPE-Bras. *Language Testing Update* 35, pp. 45-49.

- _____ (2004c) Celpe-Bras exam and teachers' perceptions of impact on teaching/assessment of Portuguese as a FL/L2. Trabalho apresentado no *Language Testing Research Colloquium*, Temecula, Califórnia, USA.
- Scaramucci, M.V.R. & Rodrigues, M.S.A. (2004) Compreensão (oral e escrita) e produção escrita no exame CELPE-Bras: Análise do desempenho de candidatos hispanofalantes. In: Simões, A.R.M., Carvalho, A.M. & Wiedemann, L. (orgs.) *Português para falantes de espanhol: artigos selecionados escritos em português e em inglês*. Campinas: Pontes, pp. 153-174.
- Shohamy, E. (1982) Predicting speaking proficiency from cloze tests: theoretical and practical considerations for tests substitution. *Applied Linguistics* 3 (2), pp. 161-169.
- _____ (1985a) *A Practical Handbook in Language Testing for the Second Language Teacher*. Tel-Aviv University, Tel-Aviv. Experimental Edition.
- _____ (1985b) Authentic language tests: where from and where to? *Language Testing* 2, pp. 48-59.
- Trites, L. & McGroarty, M. (2005) Reading to learn and reading to integrate: new tasks for reading comprehension testes? *Language testing* 22 (2). Pp. 174-210.
- Wallace, C. (1997) IELTS: global implications of curriculum and materials design. *ELT Journal* 51 (4), pp. 370-373.
- Weir, C.J. (1990) *Communicative Language Testing*. London: Prentice Hall International (UK) Ltd.
- Widdowson, H.G. (1978) *Teaching language as communication*. Oxford: Oxford University Press.
- _____ (1979) *Explorations in applied linguistics*. Oxford: Oxford University Press.
- _____ (1998) Context, Community and Authentic Language. *TESOL Quarterly* 32 (4), pp. 705-716.
- Wielewicki, H. G. (1997) *Testagem de proficiência em leitura em inglês como Língua Estrangeira numa Universidade Federal brasileira*. Dissertação de Mestrado. Santa Maria: UFSM.
- Zeidner, M. (1988) Classroom testing: the examinees' perspective. *Studies in Educational Evaluation* 14, pp. 215-233.
- Zeidner, M. & Bensoussan, M. (1988) College students' attitudes towards written versus oral tests of English as a Foreign Language. *Language Testing* 5 (1), pp. 100-114.

APÊNDICES

APÊNDICE I: Roteiro de aplicação da entrevista

1. Entrevista estruturada (perfil dos sujeitos)
 - fale sobre sua história de aprendizagem/estudo de inglês.
 - que tipos de textos em inglês você lê? Com que objetivo? (lazer, estudos, trabalho).
 - você já passou por testes de leitura em inglês antes? Fale sobre sua experiência com testes de leitura em inglês.

2. Entrevista semi-estruturada e auto-relato (percepções e descrição de comportamento não baseado em nenhum evento específico)
 - você fica ansioso quando tem que fazer um teste?
 - como geralmente se prepara para um teste de leitura em inglês?
 - o que você acha importante para se sair bem em um teste de leitura em inglês? (vocabulário, estratégias de leitura, treino no tipo de teste, conhecimento conceitual, gramática, entender significado global/local, fluência oral, pronúncia, tradução)
 - fale sobre como geralmente realiza um teste (lê primeiro o texto ou a tarefa, faz rascunho/esboço, revisa respostas).
 - que tipos de tarefas você considera mais parecidos com “exames de leitura”?
 - o que você acha das tarefas que não exigem escrita (múltipla escolha, seqüenciamento, verdadeiro/falso) e das que exigem escrita (resumo, perguntas, cloze)?
 - que tipos de tarefas você acha que são ferramentas melhores de avaliação da capacidade de leitura em inglês de alguém? Por que?
 - que tipos de tarefa você considera mais parecidos com atividades que você faria na sua vida real? No seu uso de LE?

3. Apresentação das tarefas e análise:
 - 1ª tarefa: 1 texto, questões de múltipla escolha - questionário
 - 1ª tarefa: 1 texto, resumo em português - questionário
 - 3ª tarefa: 1 texto, tarefa escrita em inglês - questionário

4. Entrevista semi-estruturada
 - Na sua opinião, qual das tarefas se parece mais com um exame de leitura? Por que?
 - Qual das tarefas você acredita que avalia melhor?
 - Qual das tarefas se aproxima mais de uma situação em que você usaria o inglês em sua vida? Explique.
 - O que você acha de tarefas que tentam simular situações do cotidiano?
 - O que você acha que é o traço mais importante, que o corretor vai mais valorizar?
 - Você considera esse traço importante também? Em que medida/por quê?
 - O fato de sempre/nunca/às vezes fazer atividades parecidas no seu cotidiano influencia em alguma coisa quando fazendo o teste?

APÊNDICE II: Teste Cloze

A tarefa abaixo faz parte da coleta de dados da pesquisa com a qual você contribuiu. Tente preencher as lacunas com as palavras que você acredita que foram apagadas do texto original. Não use dicionário e tente fazer em no máximo 45 minutos.
Muito obrigada!

THE STATE OF CHILDREN

New hope for universal health revolution for youngsters

It is generally controllable, given proper nutrition and medication. Yet half a billion children, mostly in developing countries, are affected by it each year: diarrhea. Along with (1) _____ like measles, tuberculosis, poliomyelitis, diphtheria, whooping cough and tetanus, (2) _____ claims the lives of nearly 15 million (3) _____ annually. Says James P. Grant, executive director of the United Nations Children's Fund (UNICEF): "To allow 40,000 children to (4) _____ like this every day is unconscionable in a world which has mastered the means of preventing it."

According to UNICEF's annual State of the World's Children report, released last week, things could be different. (5) _____ developing nations implemented or expanded four basic, low-cost health techniques, it suggests, they could, by the end of the century, save at least half of the (6) _____ who now (7) _____. Most of the 5 million youngsters who each year die of dehydration caused by diarrhea, for example, would (8) _____ if they were given a simple (9) _____ of clean water, glucose and salt to replenish fluids and enable their bodies to retain (10) _____. Cost per dose: 10c. The treatment, called oral rehydration therapy (ORT), was perfected in the 1960's. The systematic use of ORT has cut in half the infant death (11) _____ from diarrhea in many regions. At least 40 countries are developing comprehensive ORT, (12) _____ and UNICEF urges more.

There is also progress in the drive towards universal (13) _____ against such childhood diseases as measles, poliomyelitis and diphtheria. (14) _____ 24 months, the Brazilian government mounted six nationwide polio immunization (15) _____, each involving 400,000 volunteers, and managed to (16) _____ the incidence of the (17) _____ from 3,400 cases a year to just 26. Because of technological advances, the

(18) _____ of vaccine has dramatically decreased: measles vaccine now sells for less than 10c a dose. But further research is necessary. Some (19) _____ must be refrigerated until they are administered, which has slowed immunization programs in rural areas where (20) _____ facilities are scarce. Moreover, mustering the manpower needed to vaccinate thousands of (21) _____ in remote places remains a (22) _____.

To reduce malnutrition and infection, UNICEF promotes breastfeeding. Unlike canned or bottled formula, (23) _____ milk contains key antibodies that are passed from mother to (24) _____ during feeding. In areas where hygiene is (25) _____, babies often contract infections from bottled formula made with (26) _____ water and generally unsterile conditions. In the Philippines, a breastfeeding (27) _____ at Baguio General Hospital (28) _____ clinical infections by 87% and cut the infant death rate by 95%. Nevertheless, international aid organizations find it (29) _____ to persuade mothers in developing countries to abandon bottle feeding, which (30) _____ regard as more scientific and modern. In addition, since almost all infant malnutrition is invisible – infections and lack of nourishing food often stymie a child's growth in ways not immediately apparent – UNICEF (31) _____ that parents use charts to monitor the physical (32) _____ of their children.

The cost of such simple health (33) _____, UNICEF estimates, would be in the neighborhood of US\$6 billion annually until the turn of the Century, or one-hundredth of the world's (34) _____ on armaments each year. "It is clear", says Grant, "that a major requirement for coming to grips with critical poverty today is not money but political (35) _____. If that can be found, in order to seize the opportunities now offered, then the goal of adequate food and health for the vast majority of the world's (36) _____ need not be a dream deferred."

Time Magazine

APÊNDICE III: Materiais de coleta de dados**APÊNDICE III.I: Tarefa 1**

Study the passage and choose the best answers to the questions that follow.

QUESTIONS 1 – 10

American jazz is a conglomeration of sounds borrowed from such varied sources as American and African folk music, European classical music, and Christian gospel songs. One of the recognizable characteristics of jazz is its use of improvisation: certain parts of the music are written out and played the same way by various performers, and other improvised parts are created spontaneously during a performance and vary widely from performer to performer.

The earliest form of jazz was ragtime, lively songs or *rags* performed on the piano, and the best-known of the ragtime performers and composers was Scott Joplin. Born in 1868 to former slaves, Scott Joplin earned his living from a very early age playing the piano in bars along the Mississippi. One of his regular jobs was in the Maple Leaf Club in Sedalia, Missouri. It was there that he began writing the more than 500 compositions that he was to produce, the most famous of which was “The Maple Leaf Rag”.

1. This passage is about
 - a) jazz in general and one specific type of jazz
 - b) the various sources of jazz
 - c) the life of Scott Joplin
 - d) the major characteristics of jazz

2. The word “conglomeration” in line 1 could best be replaced by
 - a) disharmony
 - b) mixture
 - c) purity
 - d) treasure

3. In line 3, the word “improvisation” involves which of the following?
 - a) Playing the written parts of the music
 - b) Performing similarly to other musicians
 - c) Making up music while playing
 - d) Playing a varied selection of musical compositions

4. According to the passage, ragtime was
 - a) generally performed on a variety of instruments
 - b) the first type of jazz
 - c) extremely soothing and sedate
 - d) performed only at the Maple Leaf Club in Sedalia

5. Which of the following statements is true according to the passage?
 - a) Scott Joplin was a slave when he was born.
 - b) Scott Joplin's parents had been slaves before Scott was born.
 - c) Scott Joplin had formerly been a slave, but he no longer was in 1868.
 - d) Scott Joplin's parents were slaves when Scott was born.

6. The word "living" in line 8 could most easily be replaced by
 - a) money
 - b) life-style
 - c) enjoyment
 - d) health

7. The word "regular" in line 9 could best be replaced by
 - a) popular
 - b) steady
 - c) unusual
 - d) boring

8. The word "which" in line 10 refers to
 - a) regular jobs
 - b) the Maple Leaf Club
 - c) Sedalia, Missouri
 - d) 500 compositions

9. The name of Scott Joplin's most famous composition probably came from
 - a) the name of a saloon where he performed
 - b) the maple tree near his Sedalia home
 - c) the name of the town where he was born
 - d) the school where he learned to play the piano

10. The paragraph following the passage probably discusses
 - a) Sedalia, Missouri
 - b) the Maple Leaf Club
 - c) the numerous compositions of Scott Joplin
 - d) the life of Scott Joplin

APÊNDICE III.II : Tarefa 2

Yahoo! My Yahoo! Mail Search the Web Search

YAHOO! NEWS News Home - Help

[News Home](#) | [U.S.](#) | [Business](#) | [World](#) | [Entertainment](#) | [Sports](#) | [Tech](#) | [Politics](#) | [Science](#) | [Health](#) | [Most Popular](#) | [Index](#)

[Weight Loss](#) | [Sexual Health](#) | [Medications/Drugs](#) | [Parenting/Kids](#) | [Seniors/Aging](#) | [Diseases/Conditions](#)

Search: All News [Advanced](#)

Obesity, smoking add years to cells' age

REUTERS 

Tue Jun 14, 1:34 PM ET

NEW YORK (Reuters Health) - The cells of obese women or those who smoke are "older" than those of other women, suggesting that these factors accelerate aging, new research shows.

The age of a cell is indicated by the caps, or "telomeres," on the ends of chromosomes. With every successive division of the cell, the telomeres get shorter and shorter.

"The difference in telomere length between being lean and being obese corresponds to 8.8 years of aging; smoking (previous or current) corresponds on average to 4.6 years of aging; and smoking a pack per day for 40 years corresponds to 7.4 years of aging," Dr. Tim Spector, from St. Thomas' Hospital in London, and colleagues report in *The Lancet* medical journal.

Obesity and smoking are associated with oxidative stress, which, in turn, has been shown to promote telomere erosion. Therefore, the researchers hypothesized that obesity and smoking may be linked to shortened telomeres.

To investigate, Spector's team analyzed telomere length in 1122 white women between the ages of 18 and 76 years.

Telomere length fell steadily with age, as expected, the researchers found.

Further analysis revealed that the telomeres of obese women were significantly shorter than those of lean women. Smoking status was also tied to shortened telomeres and each pack-year of smoking further increased the amount of length lost.

"Our results emphasize the potential wide-ranging effects of the two most important preventable exposures in developed countries -- cigarettes and obesity," the team concludes.

SOURCE: *Lancet*, online June 14, 2005.

APÊNDICE III.III : Tarefa 3

Can Your Job Make You Sick?

Provided by [Psychology Today](#)

Try not to get worked up on the job: New research suggests that work-related stress can be as harmful as smoking.

In a study published in the British Medical Journal, Harvard University researcher Yawen Chang, Ph.D., surveyed more than 21,000 women and found that those with the most demanding jobs, little control over their work environment, and lack of support from co-workers and supervisors were more likely to suffer from stress than those more content with their jobs. They also found that women suffered physical effects: Symptoms ranged from anxiety to difficulty performing daily tasks like climbing stairs.

But in a new twist on an old issue, researchers don't suggest behavioral changes for alleviating on-the-job stress. Instead, they recommend that employees alter their work conditions--by reducing workload, increasing control over their work environment and seeking more social support.

In a related study, Tesfaye Belay, Ph.D., a microbiology researcher at Atlanta's Morehouse School of Medicine, confirmed that stress can make you physically sick. In lab studies, Belay mixed bacteria with the stress hormones epinephrine and nor-epinephrine and found that increasing the hormones' quantities caused the growth of certain bacteria--some of which can lead to health problems like pneumonia. In light of his findings, presented at an American Society for Microbiology meeting, Belay said that "if [people] take measures to help themselves relax, the risk of contracting an infection will probably be reduced."

By: Devon Madison

Originally published by Psychology Today:Nov/Dec 2000

