



UNIVERSIDADE DE BRASÍLIA
INSTITUTO DE LETRAS DEPARTAMENTO DE LINGUÍSTICA,
PORTUGUÊS E LÍNGUAS CLÁSSICAS PROGRAMA DE
PÓS-GRADUAÇÃO EM LINGUÍSTICA

**VALIDADE DE CONTEXTO DO EXAME ORAL DO IELTS:
UMA ANÁLISE SOCIOCOGNITIVA**

MARCELO VELLOSO GARCIA

Brasília - DF
2021

MARCELO VELLOSO GARCIA

**VALIDADE DE CONTEXTO DO EXAME ORAL DO IELTS:
UMA ANÁLISE SOCIOCOGNITIVA**

Dissertação de Mestrado apresentada ao curso de Pós-Graduação em Linguística do Departamento de Linguística, Português e Línguas Clássicas do Instituto de Letras da Universidade de Brasília como parte dos requisitos para a obtenção do grau de Mestre em Linguística.

Orientadora: Prof^ª Dr^ª. Ana Adelina Lôpo Ramos
Coorientadora: Prof.^a Dr^a. Gladys Plens de Quevedo Pereira de Camargo

Brasília - DF
2021

REFERÊNCIA BIBLIOGRÁFICA E CATALOGAÇÃO

Ficha catalográfica elaborada automaticamente,
com os dados fornecidos pelo(a) autor(a)

Vv Velloso Garcia, Marcelo
VALIDADE DE CONTEXTO DO EXAME ORAL DO IELTS: UMA ANÁLISE
SOCIOCOGNITIVA / Marcelo Velloso Garcia; orientador Ana
Adelina Lopo Ramos; co-orientador Gladys Plens de Quevedo
Pereira de Camargo . -- Brasília, 2021.
194 p.

Dissertação (Mestrado - Mestrado em Linguística) --
Universidade de Brasília, 2021.

1. Linguística. 2. Linguística Aplicada. 3. Avaliação. 4.
Validade. 5. Teoria Sociocultural. I. Lopo Ramos, Ana
Adelina, orient. II. de Quevedo Pereira de Camargo , Gladys
Plens , co-orient. III. Título.

BANCA EXAMINADORA

Professora Dr^a. Ana Adelina Lôpo Ramos (Orientadora)

Professora Dr^a. Gladys Plens de Quevedo Pereira de Camargo (Coorientadora)

Professora Dr^a. Viviane Bagio Furtoso (Membro externo)

Professora Dr^a. Danielle Marcelle Grannier (Membro interno)

Professora Dr^a. Patrícia Vieira Nunes Gomes (Membro externo suplente)

DEDICATÓRIA

Dedico este trabalho aos pesquisadores de língua, educadores e avaliadores do mundo.

AGRADECIMENTOS

Agradeço, primeiramente, à Mayan, minha parceira de vida, profissão e estudos. Seu auxílio incondicional, provido dia e noite, foi crucial para a conclusão do trabalho. Não há palavras que definam sua ajuda, salvo pela metáfora de que, como um oásis, aos perdidos num deserto, ela me foi abrigo.

Agradeço a Vítor, amigo de longa data, pela presença, comentários e ao crítico olhar literário que só um especialista da área seria capaz de prover ao trabalho. O auxílio quanto às inconsistências retóricas, escrita, estilo, dentre outros vários, tornaram, ao meu ver, o trabalho mais maduro, além de tê-lo aproximado de nossas bases epistemológicas. Vygotsky participava de círculos de literatura e filosofia para melhor compreender a natureza das interações; nesse sentido, sinto que o contato com Vítor me proporcionou vivências similares às do pensador russo, por ter conseguido expandir uma mesma análise por outros pontos de vista.

Ainda que posto em breves palavras, agradeço a minha orientadora, Adelina, que, como um farol no cerrado, apresentou auxílio por meio da mentoria e escrita desta dissertação de mestrado. Agradeço também à minha coorientadora, Gladys, pelo acolhimento e suporte. Por fim, a ambas, sou grato pela amizade que se firma.

Por fim, agradeço brevemente aos meus amigos Gustavo e Graciane, que foram, também, suporte no final do percurso.

EPÍGRAFE

A relação dialética entre teoria e prática requerida pelos fundamentos vigotskianos compreende prática como a mais avançada forma de teoria

James P. Lantolf e Matthew E. Poehner

RESUMO

Este trabalho visa investigar a validade de contexto do exame oral do IELTS se valendo dos pressupostos teóricos sociocognitivos vygotskianos, dos conceitos de contexto de Van Dijk (2012) e de uma adaptação do modelo de análise sociocognitivo para exames orais de Weir (2005). Investigamos o conceito de construto e sua relação com o contexto das modalidades da prova, para seguir para a investigação do construto mobilizado por meio das tarefas, da grade de avaliação e, enfim, sua validade de contexto. Trata-se de uma pesquisa qualitativo-interpretativista de natureza documental, em que exploraremos os documentos fornecidos pelo IELTS e demais filiados, a literatura específica sobre Linguística, Avaliação e Validade, a fim de construir argumentos sobre a validade de contexto desse exame. Assim sendo, em acórdância com os princípios dialéticos que regem a Teoria Sociocultural vygostkiana, tal como o Sociocognitismo linguístico e avaliativo, perpassamos a história das provas de proficiência, a Teoria Sociocultural, concepções de contexto, suas relações com diferentes linhas de pesquisa linguísticas, direta e indiretamente, principalmente voltando-as à interação e ao modelo sociocognitivo de investigação de Weir (2005) a fim de traçar contribuições a partir da tríade cultura, história e sociedade ao exame oral do IELTS. Concluimos que os elementos investigados do exame oral do IELTS têm parcial validade de contexto para fins acadêmicos, uma vez que aplica as mesmas tarefas e critérios de avaliação para dois fins diferentes. A modalidade acadêmica é, certamente, a mais afetada em termos de validade, enquanto a geral apresenta mais alinhamentos de cunho sociocognitivo.

Palavras-chave: IELTS; Avaliação; Linguística; Validade de Contexto; TSC;

ABSTRACT

This work aims to investigate the context validity of the IELTS oral exam by means of the theoretical assumptions of the Vygotskian Sociocultural Theory, the concept of context by Van Dijk (2012) and the sociocognitive model for oral exams by Weir (2005). We investigate the concept of the underlying construct in the oral exam and its relationship with the context and its modalities, then, we proceed to the investigation of the constructs mobilized through the tasks, the evaluation grade and, finally, its context validity. This research is of qualitative-interpretative nature, coming mainly from documental reports and academic papers, wherein the IELTS oral exam is explored by means of the specific literature on Linguistics, Assessment and Validity, so as to build arguments about the context validity of this exam. Therefore, in accordance with the dialectical principles that govern the Vygostkian Sociocultural Theory, as well as the linguistic and evaluative Sociocognitivist approach, we base our research on, we review the history of proficiency tests, the Sociocultural Theory, conceptions of context, their relations with varied Linguistic fields, directly and indirectly, alongside a Weir's (2005) sociocognitive research model adaptation in order to draw contributions from standpoints related to culture, history and society to the oral examination of the IELTS. We conclude that the investigated elements of the IELTS oral exam have partial context validity, as it applies the same tasks and assessment criteria for two different purposes. The academic modality is certainly the most affected in terms of context validity, whereas the general training seems to be more aligned with Sociocognitive principles.

Key-words: IELTS; Assessment; Linguistics; Context Validity; SCT;

LISTA DE ILUSTRAÇÕES

Figura 1 - Modelo de habilidade Linguístico-comunicativa	46
Figura 2 - Modelo Sociocognitivo	51
Figura 3 - Modelo Sociocognitivo Original de investigação de Validade de Contexto	53
Figura 4 - Modelo Sociocognitivo Adaptado de investigação de Validade de Contexto	53
Figura 5 - Gradação de dificuldade temática	131
Figura 6 - Alocação das tarefas em dificuldade temática e cognitiva	131
Figura 7 - Relação entre proficiência, precisão e conteúdo	147
Figura 8 - Gradação de complexidade temática	159
Figura 9 - Conteúdo das tarefas da parte 1 entrevista 1 do exame oral do IELTS	161
Figura 10 - Análise do conteúdo das tarefas da parte 1 entrevista 2 do exame oral do IELTS	162
Figura 11 - Análise do conteúdo das tarefas da parte 2 do exame oral do IELTS	163
Figura 12 - Análise do conteúdo das tarefas da parte 3 do exame oral do IELTS	164
Quadro 1 - Resumo das visões de contexto de Van Dijk	40
Quadro 2 - Resumo dos tipos de Validade do modelo sociocognitivo de Weir (2005) para exames Oraís	49
Quadro 3 - Resumo das perguntas do modelo sociocognitivo de Weir (2005) para investigação de validade de contexto	68
Quadro 4 - Panorama do exame IELTS 1	74
Quadro 5 - Panorama do exame IELTS 2	75
Quadro 6 - Transcrição e tradução das amostras do exame	76
Quadro 7 - Grade de Avaliação do IELTS com tradução	79
Quadro 8 - Relações entre os níveis do CEFR e a escala gradativa do IELTS	84

Quadro 9 - Procedimentos de pesquisa	90
Quadro 10 - Relação linguagem acadêmica e emergência no IELTS	93
Quadro 11 - Características de apresentações e seminários	96
Quadro 12 - Resumo dos elementos inerentes à apresentação e ao seminário	97
Quadro 13 - Elementos cognitivos operacionalizados	98
Quadro 14 - Resumo do formato das respostas: método de geração de dados e propósito das tarefas	105
Quadro 15 - Resumo do construto acadêmico e da análise de validade de contexto das tarefas do IELTS	110
Quadro 16 - Excerto de descritores para análise	112
Quadro 17 - Relação de horas de ensino assistido e proficiência	119
Quadro 18 - Resumo das demandas cognitivas médias por tarefa	121
Quadro 19 - Tipos de insumo	122
Quadro 20 - Análise sociocognitiva da demanda discursiva	127
Quadro 21 - Resumo das demandas discursivas de uma tarefa	128
Quadro 22 - Expectativa de produção em relação ao nível de proficiência	136
Quadro 23 - Análise da seção de “Recurso Lexical” da grade de avaliação	140
Quadro 24 - Resumo de critérios estruturais de análise 1	148
Quadro 25 - Resumo de critérios estruturais de análise 2	153
Quadro 26 - Resumo de critérios estruturais de análise 3	154
Quadro 27 - Resumo de funções informacionais em ordem de complexidade	158
Quadro 28 - Resumo de funções interacionais em ordem de complexidade	158
Quadro 29 - Resumo da administração de funções interacionais em ordem de complexidade	158
Quadro 30 - Resumo das macro e microfunções, previsibilidade da produção organizado por nível de proficiência	159
Quadro 31 - Análise da parte 1 da entrevista 1 do exame oral do IELTS	161
Quadro 32 - Análise da parte 1 da entrevista 2 do exame oral do IELTS	162
Quadro 33 - Análise da parte 2 do exame oral do IELTS	163

Quadro 34 - Análise da parte 3 do exame oral do IELTS	164
Quadro 35 - Análise das funções presentes em Ducasse e Brown (2011)	166

LISTA DE ABREVIATURAS E SIGLAS

CPE	<i>Certificate of Proficiency in English</i>
TOEFL	<i>Test of English as a Foreign Language</i>
CEFR	Common European Framework of Reference
QECR	Quadro Europeu Comum de Referência
IELTS	<i>International English Language Testing System</i>
BC	<i>British Council</i>
IDP	<i>International Development Program:</i>
CAE	<i>Cambridge Assessment English</i>
EPTB	<i>English Proficiency Test Battery</i>
ELTS	<i>English Language Testing Service</i>
TI	Tecnologia da Informação
TSC	Teoria Sociocultural

SUMÁRIO

Introdução	15
Justificativa e problemas de pesquisa	16
Objetivos da pesquisa	20
Pergunta geral de pesquisa	20
Organização da dissertação	20
Capítulo 1 - Uma breve história sobre exames	21
Capítulo 2 - Sociocognição e Sociointeração: a confluência de duas áreas revisitadas	28
2.1 O legado vygotskiano	29
2.2 Concepções sociointeracionais para aquisição de línguas - Percepções de um paradigma	33
2.3 A TSC, a interação e a Linguística	37
2.4 Contexto	39
Capítulo 3 - Sociocognição e exames - Um modelo Sociocognitivo para investigações de Validade	45
3.1 O que é construído em exames de proficiência?	45
3.2 O que é validade?	48
3.3 Definindo os critérios de análise do modelo sociocognitivo de Weir	52
3.4 Instruções, enunciados e comandos	54
3.5 Propósito	55
3.6 Formato da resposta	56
3.7 Critérios disponíveis	57
3.8 Peso	57

3.9 Ordem dos itens	58
3.10 Limitações de tempo	58
3.11 Demandas Linguísticas de uma tarefa	60
3.11.1 Modo de discurso	60
3.11.2 Características do modo de discurso: reciprocidade na interação falada	62
3.12 Canal de comunicação	63
3.13 Duração	63
3.14 Natureza da informação no texto	64
3.15 Conhecimento de conteúdo necessário	64
3.16 Insumo	66
3.16.1 Aspectos Lexicais do Insumo	66
3.16.2 Aspectos Estruturais do Insumo	67
3.16.3 Aspectos Funcionais do Insumo	67
3.17 Resumo das perguntas advindas do modelo sociocognitivo de validade de contexto	68
Capítulo 4 - Referencial Metodológico	71
4.1 A Pesquisa Qualitativa	71
4.2 Modalidades do exame	72
4.3 A pesquisa documental e o interpretativismo	86
4.4 Triangulação	87
4.5 Técnicas	88
4.6 Componentes Linguísticos de análise	89
4.7 Procedimentos	90
4.8 Considerações finais sobre o capítulo metodológico	90

Capítulo 5 - Uma análise sociocognitiva do exame oral do IELTS	91
5.1 Uma breve investigação sobre construtos acadêmicos	91
5.2 Análise das instruções, comandos, enunciados e formato da resposta	99
5.3 Análise de propósito da tarefa	106
5.4 Análise da distribuição do peso nas tarefas	110
5.5 Análise da concepção linguístico-contextual e dos descritores	111
5.6 Análise dos critérios disponíveis	114
5.7 Análise da ordem dos itens	115
5.8 Análise das tarefas	116
5.9 Análise das limitações temporais	116
5.10 Análise de insumo	122
5.11 Análise dos modos de discurso	126
5.12 Análise da extensão do insumo	128
5.13 Análise da natureza da informação	130
5.14 Análise de conhecimento do conteúdo	132
5.15 Investigação das demandas linguísticas das tarefas e da grade de avaliação	134
5.15.1 Aspectos lexicais do insumo e da grade de avaliação	134
5.15.2 Aspectos estruturais do insumo e da grade de avaliação	146
5.15.3 Aspectos funcionais do insumo e da grade de avaliação	155
5.16 Considerações finais da análise	175
Considerações finais	176
Referências	183

Introdução

O inglês, desde o início do século XIX, vem sendo utilizado como língua franca no globo terrestre. Esse *status* aponta para a crescente necessidade de conhecimento sobre o uso da língua inglesa, seja para o acesso ao entretenimento e cultura, ou para os estudos e fins profissionais. Aprender o básico dessa língua tornou-se praticamente uma obrigação. Por consequência, a necessidade de avaliar as aptidões de usuários não-nativos de países de línguas outras para fins acadêmicos ou profissionais surge, assim estabelecendo-se os exames de proficiência.

A origem de exames de proficiência de língua inglesa remonta ao ano de 1913 com o *Certificate of Proficiency in English* (CPE - Certificado de proficiência em língua inglesa¹) (WEIR, 2003). Em acordância com os princípios psicopedagógicos e linguísticos de cunho estruturalista dessa época, a prova exigia: 1) a capacidade de traduzir línguas como alemão e francês para o inglês e vice-versa; 2) o conhecimento da “alta cultura” local; 3) além de um texto dissertativo-argumentativo e um ensaio acadêmico sobre literatura inglesa; 4) o conhecimento da fonética do inglês, com um ditado oral para prática de transcrição e leitura em voz alta; 5) uma entrevista, em que se esperava um nível de escolarização próximo ao de uma pessoa nativa. Portanto, de um ponto de vista teórico, essa pessoa estaria apta a ingressar nas universidades que aceitassem o CPE.

Algumas das práticas mencionadas no parágrafo anterior, como a de tradução, ainda perduram nos corredores de universidades inglesas (WEIR, 2005). Há de se reconhecer, em contrapartida, que muito já mudou quanto à realidade da avaliação internacional de língua inglesa. As primeiras mudanças significativas foram marcadas pela chegada do TOEFL (*Test of English as a Foreign Language* - Teste de inglês como língua estrangeira) e do IELTS (*International English Language Testing System* - Sistema de Teste Internacional de Inglês, daqui em diante, IELTS) no final da década de 1970 e início da década de 1980, principalmente por implementarem elementos holísticos do uso da língua, em comparação aos modelos anteriores. (MANHATTAN, 2021).

¹ Todas as traduções presentes nesta dissertação são de nossa autoria.

Nosso objeto de estudo volta-se especificamente ao IELTS. Propomos um recorte investigativo, tratando apenas de uma seção do IELTS, o exame oral, e temporal, já que o estado atual de seu exame² oral pode vir a sofrer alterações futuras. Tais mudanças e adaptações são elementos integrantes da natureza das provas, isto é, provas de toda sorte estão sempre passando por alguma espécie de alteração pertinente à sociedade, ao tempo e à cultura em que estão inseridas. Assim sendo, empenhamo-nos para construir uma investigação de Validade de Contexto do exame de proficiência oral das modalidades *Academic* e *General Training*³ do IELTS a partir de uma adaptação do modelo sociocognitivo de análise de Weir (2005) e de princípios linguísticos sociocognitivos. A seguir, adiantamos algumas informações que podem ser úteis ao leitor, mas que ainda serão retomadas com mais profundidade adiante em nossa dissertação.

Os propósitos do IELTS são diversos. A prova pode ser utilizada para processos de imigração, notável no *Express Entry*⁴, ou para fins acadêmicos, como constatável nos editais da CAPES, os quais o listam como um dos exames aceitos, com escores que podem ir de 5 até 9, a depender do programa.

Justificativas e problemas de pesquisa

Minha trajetória profissional certamente tem peso na escrita desta dissertação. Em 2009, ao ingressar na Universidade de Brasília para o curso de Letras - Inglês, não tinha noção do que esperar em relação ao curso. Aos poucos fui desenvolvendo com mais competência as minhas habilidades com a língua com auxílio de diversos professores, até o começo de minha atuação profissional, em 2013.

Em 2016, com o país em recessão, desempregado, comecei a refletir sobre minha prática pedagógica crendo que, em princípio, eu não tinha vocação para a docência. Após tentativas em outras áreas, retornei, prodigamente, à carreira, terminei o curso de professor da Casa Thomas Jefferson, o TEFL (*Teaching English as a Foreign Language*)⁵ e procurei me certificar internacionalmente por meio da obtenção do CPE (atualmente a nomenclatura foi

² Utilizamos exame, teste e prova como sinônimos neste trabalho.

³ Referiremo-nos a elas como modalidades acadêmicas e gerais ao longo do texto.

⁴ O *Express Entry* é um dos programas de imigração do Canadá.

⁵ Ensino de inglês como língua estrangeira

alterada para C2 Proficiency) e do CELTA (*Certificate in English Language Teaching to Adults*)⁶)⁷ para ter mais prestígio profissional.

A preparação para ambos os certificados ajudou-me a amadurecer como docente e usuário da língua. Mas, notei, aos poucos, que, por mais que tivesse aprendido sobre sequenciamento didático, andaimagem, práticas pedagógicas e sensibilidade cultural, pouco debatíamos sobre a relação que existe entre o que se ensina, a operacionalização do que se deseja avaliar, e o que de fato é avaliado.

Comecei a voltar meus olhos ao tema após me tornar examinador⁸ da prova oral do IELTS em 2018. Senti-me fascinado com o universo da avaliação e achei (como ainda acho) incrível como o sistema do IELTS, como já citado acima, organizava sociocognitivamente diversos gêneros textuais de acordo com os propósitos gerais das modalidades geral e acadêmica no âmbito da leitura e escrita (IELTS, 2021).

Em contrapartida, as habilidades oral e auditiva pareciam-me menos alinhadas com seus propósitos de um ponto de vista sociocognitivo. A pergunta que, até o momento, parecia amorfa começou a tomar forma, e a inquietação que me acompanhava desde meu primeiro exame de proficiência em 2008, o Michigan ECPE (*Examination for the Certificate of Proficiency in English*) finalmente se tornou verbal: como seria possível um exame oral com diferentes propósitos possuir o mesmo arranjo para todos os casos? Descobri, ali, com o auxílio de minhas orientadoras, que o que me inquietava era referente à validade de contexto (WEIR, 2005) subjacente ao exame oral dessa prova.

Em 2019, ingressei no mestrado a fim de pesquisar o desenvolvimento de competências no âmbito cognitivo, tal como a fala internalizada e as capacidades mentais a ela relacionadas, mas, dada a minha trajetória profissional, acabei por mudar o tema de minha dissertação para uma investigação sobre a validade de contexto do exame oral do IELTS, uma vez que percebi que poderia contemplar mais elementos inerentes à minha formação como humano, pesquisador e profissional em uma só pesquisa.

Apesar de partirmos de um ponto de vista sociocognitivo, não presumimos que o exame oral do IELTS o faça. Nosso objetivo é utilizar a lente sociocognitiva de pesquisa para

⁶ Certificado de Ensino de Língua Inglesa para Adultos

⁷ O CPE (hoje C2 Proficiency) e o CELTA são, respectivamente, de acordo com o *Regulated Qualification Framework* (Matriz de referência de qualificação regulada), de nível 3 e 5.

⁸ utilizamos mediador-avaliador e avaliador como sinônimos desse termo

contribuir com observações e proposições sobre as modalidades acadêmicas e gerais do exame oral do IELTS. Pensamos que a lente sociocognitiva pode contribuir para um melhor alinhamento da validade de contexto do exame oral aos propósitos acadêmicos e gerais.

A concepção sociocognitiva de língua está ligada a outro pressuposto teórico, que é a concepção interacional, ou simbólica, que carrega em si reflexos pragmáticos do uso da língua além de suas estruturas e funções referenciais (RODRIGUES, 2005). Essa concepção prevê que o homem habita espaços interlocutivos simbólicos. O espaço interlocutivo é aquele em que, num espaço simbólico, ou seja, uma delimitação momentânea, contextual, ocorre troca de discursos e turnos dos interactantes. O espaço simbólico distingue-se, pois, do espaço físico e do espaço social. É nesse nicho que se formam e ressignificam relações interlocutivas de caráter específico e, nas palavras de Goffman (1986), teatral.

Partir dessa concepção é importante pois a concepção interacional de língua difere-se da concepção estrutural por considerar que o mundo é um contínuo em que é a própria experiência coletiva entre os interactantes que molda os conceitos que utilizamos na língua, e não o simples ato de nomear objetos, ou classificar conceitos. Ela distingue-se também da concepção referencial de linguagem que, apesar de considerar os eventuais problemas da comunicação como frutos de uma inerente plurivocalidade linguística, ainda concebia tal variação como problema. É por isso que, sob a égide da concepção interacional de linguagem, calcamos nossa pesquisa de cunho sociocognitivo.

Assim sendo, uma vez que concebemos que o *sentido* resulta da maneira como os interactantes escolhem utilizar seu repertório linguístico de acordo com o espaço simbólico que os rodeia e negociam seus discursos, não há concepção mais alinhada à natureza de nosso trabalho. Ademais, essa concepção também se estende às bases avaliativas que nos propomos a usar, como Weir (2005) e Bachman et al (1990; 1996).

Outrossim, exames orais não são tão investigados quanto os escritos pela comunidade acadêmica, tal como aponta Weir (2003; 2005). Há enorme carência de pesquisas voltadas para essa modalidade de exame, principalmente as que apontam para modelos sociocognitivos. Chapelle (1988) e Weir (2005) reconhecem que a perspectiva sociocognitiva é capaz de viabilizar pesquisas que revelem traços interacionais da performance de um candidato. Esses traços estão interligados e são influenciados pelo contexto no qual costumam ocorrer, deixando de ser uma mera operacionalização do construto por meio de elementos

linguísticos estruturais, por exemplo, e passando a ser uma operacionalização autenticamente interacional. É a perspectiva sociocognitiva que incorpora interação, construto e contexto, e é na carência de pesquisas nessa área que nossa dissertação se aloca.

Investigar a validade de contexto por meio da lente sociocognitiva, no entanto, não é o mesmo que dizer que iremos nos eximir de uma investigação que nos aponte ao menos um pouco sobre o construto que subjaz o exame oral do IELTS. No entanto, temos como objetivo principal investigar a validade de contexto do exame à luz de perspectivas sociocognitivistas.

Por fim, ao pesquisar sobre a quantidade de trabalhos referentes a exames orais e/ou aspectos de diferentes exames orais do IELTS, notamos que as fontes no Brasil são exíguas, principalmente quando comparadas ao número de pesquisas existentes voltadas para a escrita e leitura (WEIR et al, 2003). Tratando-se de uma pesquisa sobre o exame oral do IELTS no Brasil, trazemos contribuições que vão além da perspectiva sociocognitiva de análise, principalmente por conjugar-mos múltiplos pressupostos teóricos que partem de uma mesma episteme. A análise sociocognitiva discursiva abre possibilidade para diferentes utilizações dessa lente teórica, pois traz à baila visões complementares que se completam em múltiplos pontos de nossa investigação. Em outros países, pesquisas dessa sorte são mais facilmente encontradas, mas, ainda assim, não exploram o tema da mesma forma como nos propomos a fazer.

Ademais, uma vez que entendemos que testes *sempre* serão suscetíveis a ajustes e atualizações, a pesquisa sobre como melhor adequá-los à nossa realidade se encontra numa lacuna permanentemente em aberto. O estudo da validade de contexto do exame oral do IELTS faz-se necessário, pois, as habilidades orais avaliadas em um exame devem estar conectadas ao seu contexto sociolinguístico de uso, e é de suma importância investigar de que forma o contexto utilizado para operacionalizar o construto-alvo em relação a seus respectivos contextos reais são válidos em relação aos propósitos para os quais os resultados serão utilizados.

Assim sendo, apresentamos nossos objetivos de pesquisa, seguidos das perguntas de pesquisa.

Objetivos de pesquisa

- **GERAL**
- investigar a **validade de contexto** das tarefas e da grade de avaliação do exame oral do IELTS em relação aos seus propósitos sob uma perspectiva sociocognitiva;
- **ESPECÍFICOS**
- analisar as tarefas e enunciados apresentados no exame oral do IELTS
- analisar a grade de avaliação no exame oral do IELTS
- comparar **se e como** os critérios presentes na grade de avaliação estão contempladas nas tarefas e enunciados investigados
- conferir **se e como** os resultados obtidos na investigação **estão alinhados aos contextos acadêmicos e gerais** que o exame se propõe a avaliar

Perguntas gerais de pesquisa

- Considerando-se as relações entre os propósitos da prova, sua grade de avaliação e suas tarefas, como essa tríade está alinhada em relação à validade de contexto?
- Qual é a validade do contexto investigado à luz de paradigmas sociocognitivistas?
- Como é estabelecida a relação entre a grade de avaliação e as tarefas?
- O quão alinhadas, ou desalinhadas estão as modalidades da prova, a acadêmica e a geral, em relação aos seus propósitos, tarefas e grade de avaliação?

Organização da dissertação

O capítulo 1 desta dissertação apresenta um breve panorama sobre a história das provas de proficiência. O capítulo 2 apresenta um breve panorama histórico sobre as bases epistemológicas que regem nossa pesquisa e os itens modelo sociocognitivo de investigação de Van Dijk (2012) que usamos. O capítulo 3 apresenta o modelo sociocognitivo de investigação de validade de exames que usaremos (WEIR, 2005), assim como os conceitos basilares de construto e validade. O capítulo 4 mostra nosso percurso metodológico neste trabalho. Finalmente, o capítulo 5 apresenta a confluência das teorias basilares de nosso trabalho para a análise do exame em suas várias instâncias.

Capítulo 1 - Uma breve história sobre exames

Este capítulo apresenta um breve panorama dos exames de proficiência e seu desenvolvimento. O tema restringe-se majoritariamente à língua inglesa, com breves menções à história de outras línguas. O capítulo trata, destarte, do desenvolvimento de avaliações de maneira generalizada e move-se em direção à língua inglesa inicialmente voltando-se para seus exames gerais, e, sequencialmente, aos de proficiência. Apontamos as mudanças e influências históricas que alguns eventos trouxeram à cultura de avaliação estadunidense e bretã, assim como questões referentes ao uso de exames de proficiência. O capítulo, além do panorama, visa também abordar problemas referentes a exames orais de proficiência e pairar sobre problemas relativos ao nosso tema de pesquisa, isto é, exames de proficiência escritos e orais, e apresentar, ainda que brevemente, alguns pontos de nossa discussão a serem desenvolvidos mais adiante.

Spolsky (2008) aponta que as primeiras avaliações de língua documentadas em nível nacional surgiram na China, a fim de selecionar os melhores falantes de mandarim para cargos governamentais com grande compensação financeira. O sistema em questão chamava-se Exame Chinês Imperial. Neste sistema, não havia sequer uma relação constatável com o sistema educacional do país, seja em termos pedagógicos ou sistêmicos (a fim de influenciar o sistema educacional em prol de uma mudança, por exemplo); o objetivo era simplesmente selecionar os melhores usuários da língua. Países próximos como Coréia e Japão adotaram sistemas similares e mantiveram, assim como no próprio território chinês, uma tradição de testagem.

Outro teste relevante, que inclusive gera contraste com o modelo imperial chinês, chamava-se teste de Treviso⁹. O teste, segundo Spolsky (2008), focava no processo educacional, isto é, o quanto os alunos haviam retido dos componentes curriculares da instituição. O teste de Treviso indicava tão-somente o quão habilmente, sem classificação, os discentes daquela instituição utilizavam o que lhes foi ensinado dentro das premissas estabelecidas em seu currículo estudantil. O modelo de Treviso, no entanto, não durou muito, uma vez que os jesuítas, por volta do século XVII e XVIII (DE LA SALLE, 1720 apud

⁹ Treviso é uma cidade Italiana e Spolsky trata o nome do teste como uma metonímia para a universidade de Treviso e suas práticas docentes e administrativas.

Spolsky 1995) começaram a instaurar outros modelos de testagem. Na França, no modelo anteriormente jesuíta, derivado do imperial, foi separado da igreja com o auxílio da Revolução Francesa e, sob o comando de Napoleão, foi criada uma espécie de exame central, que culmina no que hoje temos, comumente, como as primeiras versões do ensino médio escolar (THÉRY, 1861 apud SPOLSKY, 2008).

A maioria dos sistemas anteriores referiam-se tanto às habilidades de oratória, quanto ao domínio de léxico e escrita, conhecimentos gerais, históricos e culturais, salvo pelo teste de Treviso, que possuía, tal como Bachman (1990) aponta como cabível, um teste baseado em seu currículo, até sua mudança estrutural de influência eclesiástica.

Uma vez que nosso objetivo é avaliar o exame e a grade de avaliação da seção oral do IELTS, cabe-nos ressaltar um dos primeiros testes de língua inglesa: o teste de Samuel Pepys para tenentes do sistema militar britânico, em que uma entrevista era aplicada aos seus candidatos a fim de substituir o sistema de escalonamento por mérito e mecenato por um sistema, tal como no imperial chinês, de classificações (SPOLSKY, 2008).

A maioria dos modelos citados anteriormente, instaurados em diversos países europeus e suas províncias (SPOLSKY, 2008), só foram alteradas nas primeiras décadas do século XX. A mudança dos campos em questão dava-se em decorrência de testes psicométricos advindos da psicologia cognitiva. A influência militar tornou os testes psicométricos ainda mais populares, particularmente após o fim da Primeira Guerra Mundial, principalmente a Avaliação Objetiva Americana (REED, 1987; YERKES, 1921 apud SPOLSKY, 2008). Com o término da Segunda Guerra Mundial e o *boom* tecnológico advindo deste embate, a Avaliação Objetiva Americana tornou-se o modelo mais popular e dominante mundo afora.

Como se davam os testes exclusivamente linguísticos, ou pelo menos, parcialmente linguísticos, ou relativos à linguagem, nesses ambientes? Segundo Spolsky (1995), a ênfase quanto ao conhecimento sobre a língua, tal como explicitado na introdução, revelava-se por meio de atividades de escrita ou tradução. Já no âmbito oral, grandes dificuldades teóricas eram encontradas, particularmente por não haver uma padronização em suas primeiras versões¹⁰, datadas do começo do século XVI. No entanto, já havia uma maneira

¹⁰ Por padronização entende-se, por exemplo, um padrão de pronúncia, ou uso. Os primeiros testes orais registrados dentro do contexto universitário de falantes nativos de língua inglesa datam do século XVI, em

estandardizada de se testar o desempenho¹¹ quanto ao uso da língua inglesa para testes de seleção de diplomatas, por exemplo, mas somente em um âmbito muito limitado, o qual Spolsky chama de *small-scale elite testing* (teste de elite em pequena escala).

Os testes voltados exclusivamente para a investigação do uso da língua inglesa de um candidato estrangeiro, seja em larga ou pequena escala, surgem na década de 1960 em Princeton e Cambridge. Ainda assim, marcados pela influência dos modelos psicométricos e atividades como tradução, escrita de redação, compreensão de textos e áudios e conhecimento da gramática, suas estruturas só foram alteradas com os estudos de Carroll (1961) e Cooper (1968) que trouxeram outras dimensões voltadas para o que hoje conhecemos como território das avaliações de línguas (SPOLSKY, 1995, 2008). As inovações foram a implementação de habilidades integradas¹² e a dimensão sociolinguística da língua.

Ainda que, aos olhos de um neófito nos estudos de avaliação, as mudanças propostas por Carroll e Copper possam parecer parcialmente satisfatórias, ainda lutava-se arduamente para garantir validade¹³ e confiabilidade¹⁴ nos testes de larga escala. Nos últimos cinquenta anos, muitos pesquisadores de avaliação, tal como Bachman, Canale, Swain, Celce-Maria Dorney, entre outros, se voltaram para uma melhoria de confiabilidade e validade por meio de esforços como estudos de competência, proficiência, procedimentos psicopedagógicos, etc. Um exemplo de esforços do tipo, investidos em exames de proficiência, podem ser vistos na história do IELTS, em que a relação validade de construto e exame eram aferidas por meio de entrevistas com docentes universitários a fim de descobrir se o teste atendia às expectativas

registros da Universidade de Cambridge, em que os candidatos deveriam utilizar o latim em uma prova oral. O problema começava a partir do momento em que não havia, por exemplo, um padrão de pronúncia estabelecido para essa língua. Logo, certo e errado tornavam-se valores subjetivos e dependentes, mais do que nunca, do padrão de pronúncia do avaliador.

¹¹ Ao utilizar desempenho, pensamos nos paradigmas chomskyanos, isto é, em oposição à competência, como um construto ideal.

¹² Trata-se de uma prática pedagógica relativamente comum nas salas de aula de hoje em dia. Durante a prática de uma habilidade, procura-se conjugá-la à outra análoga: escrita com leitura e fala com audição, por exemplo. Outras relações também são possíveis. Para mais informações sobre o assunto de um ponto de vista mais atual, principalmente no âmbito do ensino de língua inglesa, recomendamos o texto de Scrivener (2011)

¹³ Definir validade é uma tarefa árdua que depende de dois fatores, uma concepção teórica do assunto e uma concepção de verificação do assunto. Neste primeiro momento, validade e validade de construto eram definidas como “adequação no que estava sendo medido, seguida de perto pela utilidade para a comunidade de ensino.” (WEIR, 2005, p.7), isto é, “os testes seriam considerados válidos se eles se correlacionassem bem com outros testes” (SPOLSKY, 1995, p.448) **no original** -“*appropriateness in what was being measured, followed closely by utility for the teaching community.*” (WEIR, 2005, p.7) “*tests were considered valid if they correlated well with other tests*” (SPOLSKY, 1995, p.448)

¹⁴ De maneira geral, a maioria dos autores define confiabilidade como a consistência da mensuração de um dado teste ou avaliação.

dos respectivos cursos em que ele vinha sendo utilizado (SPOLSKY, 2008). Bachman (1990; 2000) foi um dos maiores responsáveis pelo avanço das pesquisas sobre validade (SPOLSKY, 2008). A pesquisa de Bachman integra implicações sociais e o uso dos resultados dos exames em seu próprio modelo de testagem. Apresentaremos a definição de validade desse pesquisador em nosso referencial teórico, adiante, uma vez que nosso foco neste capítulo é uma breve apresentação histórica.

A segunda tendência referia-se à integração de habilidades, como já mencionada anteriormente. As habilidades orais foram veementemente negligenciadas em exames de proficiência de todos os tipos. Os primeiros registros de formatos próximos aos contemporâneos são resultados de pesquisas militares. Segundo Spolsky (2008), foi Carroll quem desenvolveu um sistema de avaliação envolvendo dois ou três juízes (examinadores) para examinar o desempenho dos candidatos do exame, período de vigência da guerra fria, em que falantes com exímia habilidade oral eram necessários. Foi somente após aproximadamente vinte anos que as informações metodológicas do teste foram abertas ao público e tornaram-se modelos para avaliações voltadas ao uso da língua em órgãos governamentais e exames de proficiência.

Quando consideramos exames internacionais, o primeiro nome que surge na bibliografia especializada é o TOEFL. Inicialmente, o TOEFL era tratado como uma organização independente, mas após uma série de manobras políticas, foi tomado pela ETS, tornando-se, até os dias atuais, a sua maior fonte de rendimentos financeiros (SPOLSKY, 1995; 2008). Spolsky indica que o exame TOEFL, por vezes, encontrava-se blindado ante aos avanços teóricos propostos por órgãos educacionais e universitários. O motivo pelo qual isso ocorria era por se tratar de uma avaliação de alto valor comercial e aceitação pública, indicando que os usuários tinham mais poder de decisão dentro da estrutura do teste do que os próprios desenvolvedores em si. Uma visão menos focada em *stakeholders* (um termo utilizado para cobrir tanto usuários quanto acionistas no campo dos exames de proficiência) surge dos esforços ingleses.

Assim sendo, o desenvolvimento de testes advindos de organizações como do Sindicato Local de Exames da Universidade de Cambridge (*University of Cambridge Local Examination Board*) não contava tanto com a participação de seus usuários, mas sim de seus desenvolvedores, o que fazia com que alterações constantes permeassem a estrutura de seus

exames, lado a lado à afirmação de que se tratavam de exames tão confiáveis quanto os do TOEFL.

Cabe ressaltar, no entanto, ainda que se fizesse um teste que procurasse se atualizar com base em pesquisas, em vez de usuários, o problema maior, natural de todas as provas, permanece: a eterna imperfeição de todo e qualquer tipo de exame. A fim de amenizar a variação e garantir maior validade e confiabilidade, a implementação de grades avaliativas, ou matrizes de referência, ou correção, normalmente adequadas, ou de acordo com princípios teóricos de uma dada época que regiam o teste, eram aliadas aos exames e explicadas ou fornecidas aos seus respectivos examinadores.

Spolsky defende que a implementação de grades, ou matrizes, foi tão benéfica quanto contraditória no universo das avaliações. Os primeiros modelos criados, os de Thorndike (1910), descreviam em minúcias o que era esperado de cada nível, o que direcionava o examinador a uma alocação em termos de estágio de aprendizado e desempenho. As grades serviam como instrumento de rememoração para os avaliadores lembrarem-se de seus treinamentos, mas, ao mesmo tempo, quando breves, carregam consigo problemas como a ideia de que “a proficiência no idioma é escalonável, em vez de um conjunto de habilidades parcialmente relacionadas ao desempenho de várias funções da linguagem. [e] como resultado, eles precisam ser aceitos por consenso, em vez de validados na prática ou na teoria.”¹⁵ (SPOLSKY, 2008, p. 450)

Em oposição às grades voltadas para exames específicos, Spolsky menciona o Quadro Europeu Comum de Referência (QECR, [CEFR], COUNCIL OF EUROPE, 2001), asseverando que se trata de

Um dos desenvolvimentos mais elaborados da escala de linguagem é o Quadro Europeu Comum de Referência (Conselho da Europa, 2001), que na verdade compreende um grande número de escalas diferentes para vários tipos de conhecimento de linguagem e habilidade funcional.¹⁶ (SPOLSKY, 2008, p.450)

Mencionamos este quadro pois ele, por vezes, é citado junto dos descritores do IELTS, tal como notável no *website* do British Council (BC). No entanto, como apontam Piccardo e

¹⁵no original - language proficiency is scalable rather than a set of partially related abilities in performing various language functions. [and] As a result, they need to be accepted by consensus rather than validated in practice or theory.

¹⁶no original - One of the most elaborate developments of the language scale is the Common European Framework (Council of Europe, 2001), which in fact comprises a large number of different scales for various kinds of language knowledge and functional ability.

North (2019), o breve quadro amplamente reproduzido não passa de um exemplo de como poderíamos descrever usuários da língua, o que nos leva, erroneamente, a uma visão única de proficiência e desempenho, uma vez que o QECR não possui uma descrição para cada nível de ensino-aprendizado estática.

Ainda que esforços notáveis para a inclusão de elementos sociolinguísticos e os famosos textos de Hymes (1972), pouco ainda é visto em termos etnográficos quando o assunto volta-se à aplicação de exames de proficiência. Não há, por exemplo, seja no TOEFL ou no IELTS, uma avaliação do controle de outras variantes, bem como a funcionalidade inerentemente voltada a elas. Esse último pensamento volta-se às questões de pesquisa deste trabalho, uma vez que trata-se de um exame oral para duas modalidades diferentes, com fins diferentes em um mesmo sistema avaliativo (dada a nomenclatura do IELTS). Apesar de Bachman (1990) apontar que “autenticidade” é um conceito que beira os campos filosóficos da heurística e hermenêutica por sua inerente complexidade quanto à própria definição, Spolsky (2008) advoga, tal como nós, que avaliações não sejam feitas a partir de bases generalizáveis, o que acaba por ser, muitas vezes, objetivo final de vários tipos de exames de proficiência, segundo Fulcher e Davidson (2007).

Por mais que seja difícil adequar os testes à realidade, fazê-lo torna-se uma tarefa ética, já que há uma série de barreiras visíveis e invisíveis em decorrência da existência de testes e exames. A mais notável, em nossa opinião, é a maneira como testes vêm sendo utilizados para permitir, por exemplo, que pessoas trabalhem ou estudem em outros países. Mas, ao asseverar a primeira frase do parágrafo, não nos referimos somente ao entrave criado quanto à mobilidade global, nem afirmamos que tais entraves são desnecessários. Nosso objetivo com esse parágrafo é enfatizar os aspectos referentes ao social de cada país em que se usam estes exames e os elementos que se ligam à validade de construto, outrossim, a maneira como se aborda um determinado construto, como, por exemplo, em um exame de proficiência.

Olhar ao passado, segundo demonstra Spolsky, ajuda a corrigir as presentes avaliações.

Latham (1877) havia reclamado que as avaliações, no final do século XIX, estavam gerando objetivos educacionais muito estreitos. Avaliadores britânicos como Cyril Burt estavam conscientes quanto à influência social dessas avaliações, presumindo que eles viabilizariam crianças mais inteligentes de classes econômicas mais baixas

a tornarem-se mais móveis nas classes socioeconômicas superiores. De maneira similar, Lemann (1999), defende que o motivo principal regendo o desenvolvimento de vestibulares da parte de Chauncey era trazer uma gama maior de variedade socioeconômica ao topo em faculdades e universidades da Costa Leste e, conseqüentemente, à liderança nacional. No âmbito da avaliação de línguas, também nos voltamos preocupadamente às mazelas sociais. Edelsky et al (1983) sugere que crianças desinteressadas em testes cujos formatos eram jogos levaram os avaliadores a vê-los como inaptos quanto à sua própria proficiência linguística. (...) Shohamy (1992, 1994, 1997, 2001) elevou os níveis dessas pesquisas, detalhando as relações entre o poder de exames de línguas e controle social. Ainda recentemente, McNamara (2005) descreveu o uso de exames de línguas fora de escolas [como ferramenta] para identificar pessoas reivindicando sua ancestralidade/etnicidade ou refúgio, assim como para filtrar ou impedir a passagem de imigrantes. (SPOLSKY, 2008; p.451)

Em resumo, ainda que muito tenha mudado desde os primeiros modelos de avaliação, será o modelo do nativo intelectualmente preparado, um dos pontos a serem discutidos em nosso capítulo de análise, a base para servir de comparação com candidatos não-nativos na grade de avaliação do exame oral do IELTS? Spolsky (2008) concebe que, e nós concordamos, essa visão é retrógrada e reforça uma série de paradigmas inerentemente ligados à massificação de exames de proficiência a partir de testes ergonomicamente (um modelo que serve para todos os públicos) organizados.

A partir da contextualização do que costumava estar por trás da confecção de exames de proficiência, voltamos nosso olhar para a Teoria Sociocultural (TSC) a fim de, num breve caminhar histórico e conceitual, elencar pressupostos que nos auxiliem a investigar e contribuir para a validade de contexto do exame oral do IELTS.

Capítulo 2 - Sociocognição e Sociointeração: a confluência de duas áreas revisitadas

Neste capítulo trataremos dos princípios de Vygotsky (1986) quanto aos seus estudos de cunho sociocultural-histórica. Esses estudos são basilares para nossa investigação sobre a validade de contexto do exame oral do IELTS, uma vez que, por ser o exame, em sua parte oral, de fundamento interacional, isto é, há comunicação relativamente espontânea entre os interactantes, há indícios de instâncias co-construídas em bases cognitivas vygotskianas. Tais bases alicerçam todos os estudos de uma linguística sociointeracional. Dessa forma, Vygotsky (1986) concebia a dimensão humana como fruto da co-construção de conhecimento, isto é, os interactantes das enunciações dependem uns dos outros para adquirir, aprender e evoluir em sociedade.

Abordamos aqui os princípios básicos da teoria original vygotskiana, comentando os conceitos que mais adequadamente são relevantes à nossa pesquisa. Comentaremos, também, alguns dos avanços trazidos por outros pesquisadores à contribuição de Vygotsky, a fim de definir um de nossos pilares teóricos de análise, a Teoria Sociocultural, nossa concepção de língua e de contexto.

Tal percurso nos auxilia a traçar uma definição do que vem a ser uma língua e o que significa usá-la. Para tanto, serão considerados os trabalhos de Vygotsky (1986; 1934), Lantolf (1994; 2014; 2016; 2017) e Van Dijk (2012), por meio das relações traçadas entre a perspectiva sociocognitivo-linguística, a perspectiva sociocognitivo-contextual e, no capítulo que a este sucede, a perspectiva sociocognitivo-avaliativa. Tal aporte gera uma visão confluentemente, tripartite, em favor de nossa investigação dos instrumentos utilizados pelo exame oral do IELTS, seu contexto e sua validade de contexto. Iniciamos por Vygotsky, pedra fundamental do sociocognitivismo e do sociointeracionismo linguístico (KOCH, 2004; MORATO, 2011).

2.1 O legado vygotkiano

Esta seção visa abordar, de modo resumido, o legado do sociocognitivismo vygotkiano por meio de um panorama histórico que perpassa a formação das correntes psicológicas e linguísticas. Abordamos, então, a formação das bases teóricas psicológicas, sua influência na formação de correntes linguística e a formação de estudos vygotkianos fora da união soviética e, por fim, os elementos diferenciais nevrálgicos que contrastam o comunicativismo do sociointeracionismo. Começamos por um breve panorama histórico.

É inegável que a corrente positivista científica foi intensamente responsável por promover boa parte dos estudos linguísticos do século XIX até meados do século XX. Ferdinand de Saussure (1916) e Leonard Bloomfield (1933) são grandes nomes desses linguistas que desenvolveram estudos estruturalistas basilares para a formação das concepções de várias das correntes linguísticas da modernidade.

A corrente positivista, no entanto, não começa e nem termina na linguística. Essa também afetou outras correntes teóricas, tais como as Ciências Sociais, o Direito, a Economia e a Psicologia. Essa última, em particular, fundada inicialmente a partir de estudos de Ivan Pavlov (1927), Wladimir Bekhterev (1917) e, posteriormente, Burrhus Skinner (1953; 1957), o mais difundido para fins educacionais, visava analisar a maneira que as reações humanas poderiam ser moldadas sob condições específicas a partir do estudo de estímulos e respostas, com desdobramentos significativos para a compreensão da aquisição e de aprendizagem, resultando no surgimento da abordagem estruturalista de ensino tanto em letramento de L1, como de L2/LE.

Apesar de as correntes de pesquisa ligadas ao positivismo e ao comportamentalismo (behaviorismo) ainda exercerem grande influência nos tempos atuais, elas, em sua época, não passaram sem críticas. Estudiosos como K. N. Kornilov (1924; 1930) e Jean Piaget (1923; 1924; 1926) a criticavam e direcionaram suas próprias pesquisas à importância de estudos psicológicos que considerassem o ambiente biossocial, isto é, não só a organização de uma dada sociedade, mas também a fauna, flora e cultura local de seus sujeitos para que uma análise mais fidedigna do comportamento e do funcionamento da mente daqueles colaboradores pudesse ser feita.

Marca-se, então, o começo de estudos psicológicos dialéticos (FREITAS, 2003) entre

desenvolvimento da mente e comportamento, oposto ao construto de *tábula rasa*, tipicamente comportamentalista¹⁷. No entanto, as primeiras tentativas acabaram por ser protótipos do que uma psicologia dialética deveria ser, pois interpretavam somente

etapas gerais de pensamento ou a habilidade geral de classificar, pensar logicamente e se lembrar. Esperava-se que alguns indivíduos (ou grupos) estivessem em etapas “superiores” ou tivessem habilidades de classificação, lógicas e de memória melhores – em geral – que outras pessoas. A pesquisa intercultural era utilizada para examinar, sob circunstâncias amplamente variáveis, quais fatores ambientais gerariam maior “competência” (ROGOFF, 2005, p.42)

Ao unificar premissas comportamentais à formação de consciência através de uma estrutura sociocultural, a premissa anterior acabara por comparar, contrastar e categorizar, diferentes sociedades por meio de habilidades silogísticas e resolução de problemas matemáticos, testes de conhecimento linguísticos, dentre outras categorias de cunho estruturalista a fim de verificar os diversos desenvolvimentos cognitivos presentes nas regiões pesquisadas. Tal tipo de avaliação é problemática por procurar medir diferentes desenvolvimentos sociais a partir de um modelo individual, idiossincrático. Críticas, como as de Michael Cole (1996), corroboram com a afirmativa anterior de que os resultados lidavam somente com categorias formais do conhecimento reconhecidas pelas instituições financiadoras das pesquisas. Consequentemente, grupos sociais que não apresentavam as categorias em questão eram classificados como menos desenvolvidos cognitivamente.

Cole (1996) constata que a falta de familiaridade com uma habilidade formal somente pode ser julgada como (ir)relevante por um determinado grupo social. Falhar em um teste como o anterior não excluiria a possibilidade de desenvolvimento cognitivo em áreas diferentes, pois, na formação da mente daquele sujeito, pode haver foco em outras habilidades de maior relevância para a sua sociedade. Isso implica que, por mais que os testes levassem o ambiente biossocial dos sujeitos da pesquisa em conta, eles não consideravam as idiossincrasias presentes naquele ambiente biossocial. A percepção da necessidade de englobar tais elementos fortaleceram a ideia de que a

(...) cognição é uma competência geral que caracteriza os indivíduos em diferentes situações, [análises monolíticas evidenciavam] desequilíbrio no desempenho [que] era desconcertante. Para tentar resolver a diferença na “habilidade” aparente em

¹⁷ “Expressão latina tornada célebre por Locke e Leibniz a propósito do debate em torno do *inato* e do *adquirido*. Para Locke, o espírito humano é, desde seu nascimento, como uma *tabula rasa*, adquirindo todos os seus conhecimentos pela experiência.” (JAPIASSÚ; MARCONDES, 2008)

situações distintas, os pesquisadores tentaram primeiramente tornar mais familiares o conteúdo e o formato dos testes, para encontrar parâmetros mais “verdadeiros” de competências subjacentes. Os pesquisadores também tentaram subdividir a competência em “domínios” menores, tais como o conhecimento biológico e o conhecimento físico, ou habilidades verbais e não-verbais, de forma que as discrepâncias entre situações não fossem tão grandes (...)” (ROGOFF, 2005, p.43)

Vygotsky (1986), fundador da linha de pesquisa sociocultural-histórica, concebe, como elucidada Lantolf et al (2014; 2016), que as funções sociais eram co-construídas durante a formação e desenvolvimento da língua. Caso não fosse de tal forma, o processo de aquisição da primeira língua não seria diferente daquele que envolve uma segunda língua, ou outra qual fosse. Isto é, ao aprender uma língua, o processo de internalização da fala é o último e mais complexo a ocorrer e, conseqüentemente, o início das reflexões metacognitivas mediadas pela língua que um aprendiz realiza. Definem-se funções mentais superiores a partir dos processos de desenvolvimento cognitivo que, nas palavras de Freitas (2003), são:

a) as funções mentais superiores não podem ser reduzidas a processos elementares, existindo diferentes níveis de funcionamento psicológico com características próprias e irreduzíveis; b) o pensamento e a consciência não são uma emanção de características estruturais e funcionais internas, mas são, ao contrário, fortemente influenciadas por atividades externas e objetivas, realizadas num ambiente social. Essa é a ideia da gênese social do pensamento; c) atividade é definida como uma unidade de análise, integrando as características sócio-interativas e individuais-cognitivas das condutas; d) a cooperação social na atividade se realiza por meio de instrumentos ou signos dentre os quais os signos verbais desempenham papel primordial. (...) é pela internalização progressiva desses instrumentos de cooperação que se constrói o pensamento consciente, que transforma e regula as outras funções psíquicas. (FREITAS; 2003, p.88-89)

A postulação de que a fala social precede a egocêntrica foi uma das maiores contribuições do interacionismo vygotskiano. Lev Vygotsky (1986; 1934) foi um dos responsáveis por salientar que o conhecimento depende de uma tríade indissociável: a relação entre cultura, história e sociedade, linha a qual ele chama de sociocultural-histórica. De acordo com esse psicólogo russo, a formação de conhecimento¹⁸ é disforme, revolta, sinuosa e dependente de práticas sociais¹⁹, isto é: o eu, o outro, onde e quando vivemos.

O princípio fundamental da teoria vygotskiana traduz-se, então, como a relação dialética e interdependente existente entre o que é externo e o que é interno ao ser humano e

¹⁸ Neste caso enxergamos competência e conhecimento da mesma forma.

¹⁹ “o sujeito não se constituía a partir de fenômenos internos e nem se reduzia a um mero reflexo passivo do meio. Para ele, o sujeito se constituía na relação. A consciência também não era apenas a fonte dos signos, mas o resultado [do confronto] deles.” (FREITAS, 2003, p.110)

de que modo isso se estabelece. Revolucionando a maneira de o ser humano tornar-se cognitivo, opondo-se aos preceitos deterministas, previsíveis e comportamentalistas, de um lado, e inatista, de outro, o estudioso russo observa quão importante é a inserção social do ser humano para organização de seu conhecimento, do mesmo modo de como o é na interferência desse no mundo empírico. Assim sendo, é de excelência a premissa do autor, segundo a qual

Todas as funções no desenvolvimento da criança aparecem duas vezes: primeiro, no nível social, e depois, no nível individual; primeiro, entre pessoas (interpsicológica), e, depois, no interior (intrapsicológica). (VYGOTSKY, 1984, p.64).

Dessa afirmação, decorrem conceitos pilares da teoria, os quais estão contemplados, de modo total ou em parte, em diversas áreas do conhecimentos, entre as quais a linguística cognitiva de cunho sociointeracional se encontra situada. Interação, mediação, fala internalizada (social), andaime coletivo e zona de desenvolvimento real e zona de desenvolvimento proximal estão contemplados em estudos de linguagem, muitos deles, com reflexos em práticas didático-pedagógicas. (JOHNSON, 2003).

Para a pesquisa aqui proposta, interessa-nos de perto o conceito de interação e de mediação por se tratar de análise sobre a parte interacional do exame IELTS, em que a língua deve cumprir papel de mediadora simbólica entre candidato e as práticas de contexto, e o contexto simulado por meio do mediador, manifesto nas tarefas no exame oral propostas.

2.2 Concepções sociointeracionais para aquisição de línguas - Percepções de um paradigma

A partir do legado vygotskiano, comentaremos, brevemente, de que modo a Teoria Sociocultural (LANTOLF et al, 1994; 2014)²⁰ influenciou as pedagogias sociointeracionistas, e sua influência nos exames de proficiência .

Primeiramente, é importante ressaltar o que Lantolf e Poehner (2014) explicam sobre a Teoria Sociocultural (doravante TSC) vygotskiana: ela não é e nem pode ser vista como uma teoria somente sociolinguística, mas sobretudo linguística cognitiva, portanto, psicolinguística de cunho sociocultural-histórica. O que essa teoria objetiva investigar é a psicologia humana englobando, em que se inclui o desenvolvimento da aquisição de línguas como unidade dialética biologicamente endossada pelo funcionamento do cérebro juntamente às formas de mediação mentais geradas pela interação sociocultural.

Em vista disso, partimos, como os autores referenciados, de uma perspectiva sociocognitiva para investigação de contexto, aportada nos estudos de Van Dijk (2012), que propõe esse importante aspecto sociocognitivo como um construto mental, resultado da soma do “lugar” cultural (mais geral) com o “lugar” assim situacional (específico) (HALLIDAY, 1976), sobre o que passaremos a comentar na seção 2.4.

O segundo pressuposto que será considerado para nossas bases linguísticas investigativas encontra-se aportado em Wertsch (1985), com diálogo estreito com a TSC, é o pressuposto subjacente à Unidade de Análise²¹ (WERTSCH, 1985), cuja unidade microcosmológica é definida de maneira concreta, isto é, uma ferramenta mediadora utilizada para investigar o processo de formação das funções mentais superiores, pois

para Vygotsky [o desenvolvimento] da língua é uma função mental complexa que se desenvolve graças à mediação e co-construção de significado nas comunicações interpessoais²² (PICCARDO, NORTH, 2019, p. 73).

²⁰ A abordagem sociointeracional em sentido estrito nunca existiu no ensino de língua inglesa, ao menos, não com este nome. As primeiras correntes que parecem abordar os ideias vygotskianos são resultantes do documento QECR (2001) e que, recentemente, parecem ter se concretizado na chamada Perspectiva Acional (PICCARDO; NORTH, 2019). No entanto, levando em conta que o sociointeracionismo é mais do que uma corrente pedagógica, pode-se notar vários elementos teóricos advindos da teoria original em muitas abordagens de ensino diferentes desde a década de 80. Ramos (2014), a concebe como existente no ensino de PBSL.

²¹ Tradução nossa, no original - Unit of Analysis (LANTOLF, 1994)

²² No original - for Vygotsky language is a complex mental function that develops thanks to the cultural mediation and co-construction of meaning in interpersonal communication

A Teoria da Atividade e a Unidade de Análise são consideradas parte das teorias vygotskianas (LANTOLF et al, 1994; 2014) de forma controversa dada a sua natureza materialista. Materialista, porque Vygotsky acreditava que os signos, concretos ou abstratos que compõem e acompanham a linguagem, que também é concebida como signo e ferramenta, simultaneamente de forma concreta e abstrata, eram mediadores de toda qualquer atividade psicossocial, histórica e cultural. O processo de formação das relações dialéticas entre o plano externo (social), o interno (mental) e suas metatransformações (JOHNSON, 2004) eram mais relevantes do que a descrição dos materiais mediadores, uma vez que estes são meramente reflexos de um tempo. Em outras palavras, a relação existente entre ferramenta, função e signo, era o que mais importava a Lev Vygotsky.

Diferentemente dos postulados originalmente simbólicos de Vygotsky (1986), a Teoria da Atividade e Unidade de Análise são alvo de controvérsia como parte do sociointeracionismo vygostkiano por considerarem que seria impossível investigar ferramentas e a maneira como são utilizadas de maneira apropriada sem entender, primeiramente, a comunidade ou cultura em que são usadas e ressaltarem o aspecto materialista como representação fundamental do mediador para a mudança social e mental porque “todo comportamento humano é organizado e controlado por artefatos materiais (ex: concreto) e simbólicos (ex: semióticos)” (SWAIN et al., 2015)²³. Para desambiguar esse impasse teórico, Lantolf (1994, 2014, 2017) e Johnson (2004) apontam para possível adaptação da TSC original a partir da consideração de que ferramentas também pertencem a uma realidade sociohistórica-cultural, ainda que sua utilização e forma possam ser alteradas conforme a época em que são utilizadas.

Assim, Piccardo e North (2019), Lantolf e Poehner (2014) e Miller (2011) identificam e classificam a mediação em três níveis:

- (a) Primeira ordem – metacognitiva – a qual origina a comunicação interpessoal e funções regulatórias;
- (b) Segunda ordem – cognitiva – a qual relaciona-se a ferramentas culturalmente construídas que nos ajudam a definir conceitos e conhecimento;

²³ No original - [a]ll human behaviour is organised and controlled by material (i.e. concrete) and symbolic (i.e. semiotic) artefacts

(c) Terceira ordem – social – a qual relaciona-se com o nível macro, isto é, as instituições e a sociedade que definem as mediações de primeira e segunda ordem.

Wertsch (1985) sugere que a morte prematura de Vygotsky talvez explique a falta de estudos envolvendo qualquer espécie de mediador material com potencial operacional para os conceitos postulados por esse estudioso; essa afirmativa faz com que a expansão da Teoria da Atividade, ou seja, a concepção que considera as conexões criadas por meio de ferramentas parece estar intimamente ligada com a concepção de formação de culturas, sociedades e práticas sociais, aspecto o qual se esclarece em Bakhtin (2017; 2019), Tomasello (2003), Freitas (2003), Johnson (2004) e Ramos (2007).

A história por trás dos conceitos teóricos sociointeracionais serve de aporte dialético para a maneira como concebemos língua. Antes de apresentar o conceito de língua aqui adotado, cabem alguns esclarecimentos comuns, principalmente, inerentes à percepção de alguns docentes e acadêmicos sobre a concepção linguística, além da psicopedagógica da TSC (LANTOLF, POEHNER, 2014). Não é incomum acharmos definições do legado vygotskiano longe da complexidade da TSC, voltadas aos princípios psicopedagógicos da teoria, somente. Por vezes, autores como Ellis (2003) atribuem a concepção de língua em uma perspectiva social à abordagem comunicativa, até pela forte presença do Comunicativismo no ensino de inglês como língua estrangeira, e pela maneira como se concebe co-construção e função dentro dos postulados funcionalistas de Halliday (1976), aparentemente alinhados à definição de mediação e co-construção de Vygotsky. Mas, na visão vigotskiana, há avanços em trazer a face funcional da linguagem à perspectiva de um sujeito, em sua cognição, cujo desenvolvimento é fruto de relações inter e intraindividuais, evidenciando tanto o lado objetivo (sociocultural) como o subjetivo (cognitivo) . Assim sendo, a ação de interagir vai além do ato de comunicar, embora aquela contemple esta atividade. Para isso, definimos, a partir de Ramos (2014) que nos diz

a interação, considerando os contextos socioculturais dos usos linguísticos.(...) vai além do ato de comunicar (...) é agir na aquisição e na construção do conhecimento e do desenvolvimento de habilidades (...) nas práticas interativas que os indivíduos adquirem recursos (...), como: vocabulário, sintaxe, conhecimento sobre mudança de turnos, aspectos retóricos e habilidades.” (JOHNSON apud YOUNG, HALL, 2003, p. 97) A interação (...) É vista como um processo de co-construção, definido como “um conjunto de criação de dada forma, interpretação, postura, ação, atividade, identidade, instituição, habilidade, ideologia, emoção ou outra realidade de significação cultural (RAMOS, 2014, p.148-151)

Essa maneira de interpretar o vínculo dialético estabelecido entre ferramentas, culturas e sociedades nos direciona às propostas de Bakhtin, contemporâneo de Vygotsky, quando ele assevera a necessidade de conceber língua, cultura, sujeitos e práticas sociais como unidade, independente das diferenças de registro ou da forma, em oposição, por exemplo, à dicotomia saussuriana de língua e fala. Bakhtin procura demonstrar, por meio de seu cronotopo, as relações dialógicas entre interlocutores e seus assuntos (chamados de heróis - aqueles que entendem completamente as intenções do enunciador original), o reflexo sociotemporal que seus discursos carregam, assim como a impossibilidade de análises estruturais da composição de obras, traço marcante de vários de seus intelectuais contemporâneos russos (BAKHTIN, 2019). As manifestações dialéticas são definidas por Bakhtin como gêneros do discurso, estando o discurso presente tanto num plano mental e social quanto material²⁴. O discurso é representante da concretização dos textos²⁵ e da heterogeneidade dos gêneros do discurso.

Para os estudos da TSC, da interação e da filosofia da linguagem, as proposições dialéticas de Bakhtin são, também, úteis. O dialogismo, para o autor russo, está em todas as dimensões da linguagem. Para ele, ouvir uma música, ou ler um texto, significa entrar em contato com um enunciado que dialoga com seu destinatário, gerando assim, constante intertextualidade e contato social. A língua, nesse contexto, é um contínuo polifônico, composto por textos e contexto, no tempo e espaço, que concretiza relações sociais por meio de vários signos, os vários gêneros do discurso, que Vygotsky não considera explicitamente, em suas obras.

Na próxima seção, ocupamo-nos de abordar a TSC e de definir interação e contexto a partir de perspectivas linguísticas de cunho sociocognitivo. Logo, a principal função da seção que segue é abordar a importância do caráter da interação e da linguagem com seus contextos de emergência.

²⁴ Bakhtin (2017;2019;) organiza os gêneros do discurso em primários e secundários. O primário são definidos em situações cotidianas; os secundários são definidos por atividades enunciativas complexas e especializadas. Uma distinção simples pode ser feita entre um bilhete e um romance. (RAMOS, 2007) No entanto, estes gêneros não são categorias imutáveis e podem se adaptar de acordo com o momento histórico em que são empregados, isto é, o tempo dirá quais são primários e secundários uma vez que é a época em que eles ocorrem que define seu nível de complexidade.

²⁵ Lemos texto como sinônimo de enunciado, fala, texto escrito, visual, etc.

2.3 A TSC, a interação e a Linguística

Hymes (1972) e a competência comunicativa, assim como John Gumperz (1972; 1982) e suas *Estratégias do Discurso* são um marco para uma visão linguística sociointeracional. Os autores, originalmente, concebiam-se como sociolinguistas e teciam reflexões sobre a maneira como as correntes linguísticas vigentes tratavam e concebiam língua naquele momento histórico.

Hymes (1972), responsável pela concepção de competência comunicativa, afirma que ela “é integral aos valores das atitudes e motivações ligados à língua, suas características e propósitos, integral à competência em favor, e em ações que se mobilizam, da inter-relação da língua com outros códigos de conduta comunicativos²⁶” (HYMES, 1974, p.277-8).

Os princípios do sociocognitivism, de modo direto ou indireto, chegam à linguística após a primeira metade do século XX, instigando estudiosos a observarem a língua nas instâncias das situações enunciativas, nas instâncias interacionais.

Gumperz (1982), um dos ícones dos estudos sociointeracionais, demonstra que “a escolha de um indivíduo do estilo de fala tem valor simbólico e consequências interpretativas que não podem ser explicadas simplesmente por meio da correlação de ocorrência de variantes linguísticas determinadas por categorias sociais e contextuais independentes²⁷” (GUMPERZ, 1982, p.vii), mas sobretudo, pelas negociações estabelecidas pelos atores sociais nas interações. Vale notar que Gumperz inaugura esta afirmativa em seu trabalho de 1982, mas já vinha pesquisando diferentes maneiras de concretizar tal concepção heurística e epistemológica quanto à formação da língua e da linguagem como construtos interacionais em trabalhos anteriores²⁸.

Gumperz foi influenciado por pesquisas sociológicas, notadamente, por Erwing Goffman, um sociólogo que se propôs a estudar as relações humanas em um nível

²⁶no original - is integral with attitudes values and motivations concerning the language, its features and uses, and integral with competence for, and attitude towards, the interrelation of language with the other codes of communicative conduct

²⁷ no original - an individual's choice of speech style has symbolic value and interpretive consequences that cannot be explained simply by correlating the incidence of linguistic variants with independently determined social and contextual categories.

²⁸ Ver *The Ethnography of Communication* (1972) e *Language in Social Groups* (1971) para mais detalhes sobre o trabalho de Gumperz..

microsociológico e teatral, que destaca-se entre os interacionistas como um dos mais citados. Enquanto outras teorias sociais preocupavam-se em descrever o funcionamento da sociedade de um ponto de vista macrossocial, Goffman ocupava-se de analisar enquadres (*frames*), ou seja, o desempenho comportamental dos sujeitos de uma determinada sociedade em situações de toda sorte em seus cotidianos. Para viabilizar seus ensaios de cunho sociointeracional, Goffman propõe sua teoria apresentando o conceito de *Jogo*, descrito em *Rituais de Interação* ([1967]; 2011), em que postula como a interação social é regida por ações premeditadas e socialmente estabelecidas; em seu *magnum opus*, *Frame Analysis* (1986), ele atribui um valor teatral aos agentes, bem como papéis sociais intrínsecos a cada um dos enquadres analisados, a fim de viabilizar interpretações e descrições inerentemente dependentes de seus ambientes e práticas sociointeracionais.

Piccardo e North (2019) asseveram que os trabalhos de Hymes, Goffman, (e nós adicionamos Gumperz à lista), expandiram o campo da Linguística Aplicada. As considerações anteriores, apesar de referirem-se a pesquisas linguísticas sem fins pedagógicos, impactaram as concepções de língua vigente e trouxeram à tona, ainda que indiretamente, vários ideais vygotkianos.

Dentre os notáveis escritores da área, podemos ver que, no âmbito sociocognitivo, educacional e interacional, destaca-se Johnson (2004), que sugere como podemos conceber competência por um prisma sociointeracional em contexto de segunda língua.

A interação neste modelo é vista não como uma questão cognitiva, mas como uma questão social. Ou seja, a interação não é vista de forma abstrata, como um processo de combinação de diferentes competências que ocorre dentro da mente do indivíduo, mas é vista como uma interação face a face entre todos os participantes em práticas interativas. Assim, a interação assume uma forma diferente - a forma da comunicação humana entre as pessoas em ambientes socioculturais da vida real. (JOHNSON, 2004, p.97)²⁹

Entendendo interação como central para a aquisição de habilidades e línguas, alinhando às teorias cronotópicas de Bakhtin e à TSC, de Vygotsky e, por fim, as proposições linguísticas que influenciaram a Linguística Aplicada, concebemos língua como um conceito

²⁹ no original - Interaction in this model is viewed not as a cognitive issue but as a social issue. That is, interaction is not viewed abstractly, as a process of combining different competencies that takes place inside the individual's mind, but is viewed as face-to-face interaction among all participants in interactive practices. Thus, interaction takes on a different form—the form of human communication among people in real-life sociocultural settings.

que se apoia na tríade história, sociedade e cultura, e que se realiza, evolui, negocia-se e altera nas relações intersociais que são realizadas por meio de si.

Ao fechar a concepção de língua e interação, é preciso trazer à tona mais um conceito central às discussões que aqui propomos: o de contexto. As informações sobre contexto, a seguir, se relacionam de algum modo ao de Weir (2005), que, embora o considere sob ótica dos campos referentes à validade, implica a noção de cultura, de situação, portanto de construto mental do candidato ao exames de proficiência, portanto, assentados em fundamentos sociocognitivos já descritos. A principal finalidade da seção seguinte é, portanto, discorrer sobre a noção de contexto como construto sociocognitivo.

2.4 Contexto

Uma vez que nosso trabalho tem como objetivo investigar a validade de contexto do exame oral do IELTS, há necessidade de subsídios para que possamos definir, primariamente, o que é contexto para, então, adentrar os estudos de validade de contexto por uma perspectiva também sociocognitiva.

Para nossa investigação, diferentes visões confluem, sendo a primeira, aquela cuja tônica contempla a ideia de modelo mental, que adotaremos, por trazer em si o aspecto da sociocognição. Estamos falando do conceito de Van Dijk (2012), pesquisador holandês, que é um notável linguista e analista do discurso. Muito de sua obra volta-se para o estudo e crítica das relações de poder que tacitamente se instauram na sociedade, o que não se relaciona diretamente com essa dissertação. Contudo, há de se notar que, além de ser um dos poucos pesquisadores que adiciona a base sociocognitiva a sua linha de análise do discurso crítica, ele também apresenta contribuições sobre o desenvolvimento sociocognitivo da construção de práticas discursivas em diversos contextos sociais, o que nos é particularmente útil como aporte sociocognitivo-linguístico para a investigação da validade de contexto do exame oral do IELTS.

Para Van Dijk (2012), contexto é um modelo mental, subjetivo, cujos participantes de uma determinada situação, seja ela interacional, social ou comunicativa, acionam ao adentrá-la. Suas categorias são inúmeras e servem de guia para nossa investigação dos

contextos subjacentes à prova, assim como para a validade de contexto da prova oral do IELTS. A seguir, apresentamos um quadro resumo de cada uma das visões que consideramos relevantes para a investigação do alinhamento da prova oral com seus propósitos gerais e acadêmicos. O modelo sociocognitivo de análise da produção de discursos em relação ao contexto é, certamente, uma grande contribuição, mas dada a natureza breve da dissertação atemo-nos aos que utilizamos em nossa análise. Vale ressaltar que Van Dijk (2012) aborda muito o termo *construto*, que, por hora, definiremos como habilidade psicológica. Desenvolvemos o termo *construto* mais a fundo no próximo capítulo.

Quadro 1 - Resumo das visões de contexto de Van Dijk

Tipo e manifestações de(o) contexto	Definição
Construto subjetivo dos participantes	Um, ou vários, construtos dos participantes, ou definições subjetivas das situações interacionais ou comunicativas. O autor não desconsidera estruturas sociais e políticas e sua influência na produção de discurso. Situações sociais só conseguem influenciar o discurso através das interpretações que delas fazem os participantes. As situações sociais, em geral, são construídas socialmente; elas só conseguem influenciar a conduta humana enquanto forem parte de uma estrutura social que as aceite.
Modelo mental	Os construtos subjetivos dos participantes podem ser explicados por modelos mentais dependentes de <i>modelos de contextos</i> . Esses modelos representam as características relevantes do ambiente comunicativo presente na memória experiencial do sujeito, e controlam o desenvolvimento cognitivo da produção e compreensão do discurso.
Modelo esquemático	Os modelos mentais e seus contextos fazem parte de esquemas compartilhados, convencionais e dotados de base cultural que auxiliam rápida interpretação de eventos sociohistóricoculturais. Os participantes não conseguiriam entender, representar e atualizar situações sociais sem tal categoria. Tempo, lugar, participantes, suas diferentes identidades, e/ou papéis interacionais, ações, propósitos e conhecimento compõem esse modelo esquemático. Identificar a organização de tais contextos depende de teorias empíricas. Cada situação, ou inter/intra-situação gera novos esquemas. Gênero e/ou ocupação podem ser mais ou menos relevantes ou proeminentes em diferentes situações, assim como variáveis num mesmo contexto.
Fundamento social	Os contextos são únicos e subjetivamente dependentes das situações comunicativas; sua estrutura e construção têm base social. Um exemplo de manifestação parte das <i>cognições sociais compartilhadas</i> (conhecimentos, atitudes, ideologias, gramática, regras, normas e valores) de uma comunidade discursiva, o que também acontece com categorias esquemáticas. Essas últimas definem as estruturas dos contextos. Os contextos, então, também têm uma dimensão intersubjetiva, que é condição primeira para a comunicação e a interação. Contudo, ‘fatos’ podem ser interpretados de uma maneira subjetiva peculiar, diferentemente do que foi estabelecido originalmente naquele fundamento social. Em resumo, contextos são pessoais e sociais - o que se aplica aos discursos que eles controlam. Cognições sociais, recursos linguísticos e comunicativos, podem igualmente ser definidos como parte ou base de um dado contexto, uma vez que esses recursos são fragmentários, tal qual os contextos, que também o são.

Planejamento	Os participantes, em teoria, conhecem previamente e, por isso, planejam o seu discurso perante uma situação comunicativa. Assim, eles observam com quem irão se comunicar, quando, onde e porque. As manifestações mais notáveis valem para os modos de comunicação escritos e normais, mas também se aplicam à interação espontânea. Em resumo, eventos comunicativos são frequentemente encaixados em eventos sociais. As representações ou experiências manifestam-se junto de interações em andamento. O ato de antecipar as atitudes próprias a um contexto ocorre porque é a partir de esquemas e categorias gerais de contextos culturalmente partilhados que se constroem contextos específicos. Assim sendo, as pessoas utilizam-se de suas lembranças para lidar com eventos comunicativos relacionados a outros passados. Presume-se, então, que é possível ‘aprender’ os gêneros discursivos, assim como tipos de contextos.
Contexto versus Texto	Contextos não podem ser resumidos ao texto ou à fala. Eles são “ <i>não textos</i> ”. Os modelos de contexto permanecem em grande medida pressupostos. Eles influenciam a fala e a produção do texto de maneiras indiretas explicitadas nas circunstâncias específicas em que são produzidos. É fato, por exemplo, que um palestrante famoso não precisa dizer quem ele é. O público já saberia disso pois seria parte do contexto. Isso indica que contextos são sinalizados ou indiciados e, muitas vezes, suas propriedades precisam ser inferidas a partir de estruturas interacionais e variações do discurso. Podemos ver um exemplo nas conversas do dia a dia, assim como em muitos tipos de fala institucional, em que podem acontecer referências a outros textos e conversas prévios. Os discursos da mídia relacionam-se de maneira múltipla a vários ‘discursos-fontes’. Em resumo, a intertextualidade discursiva pode ser uma condição importante para que o discurso seja significativo e apropriado.
Macrocontextos e microcontextos	Os modelos de contextos podem representar situações sociais ou comunicativas em vários níveis de <i>generalidade</i> ou gradação. Por um lado, modelos podem representar, num nível micro, interações momentaneamente situadas, e, por outro lado, situações históricas ou sociais, assim sendo a estrutura social num nível macro. Esses níveis podem variar no interior do mesmo evento comunicativo e podem ou não ser indiciados como tais durante o texto ou a fala.
Manifestações semânticas e pragmáticas	O discurso e suas propriedades podem expressar propriedades dos contextos de muitas maneiras. <i>Expressões indiciais</i> ou <i>dêiticas</i> são exemplos. Contudo, a semântica e a pragmática dessas expressões devem ser pensadas. A descrição da <i>referência</i> a elementos da situação em processo, como o tempo presente, falante ou destinatário da mensagem, em expressões como <i>agora</i> , <i>eu</i> e <i>você</i> , fazem parte de um tratamento <i>semântico</i> do discurso. O tratamento <i>pragmático</i> não diz respeito ao referencial, mas à adequação das expressões à situação em curso. Os pronomes <i>tu</i> e <i>vous</i> em francês se equivalem semanticamente, mas diferem pragmaticamente evidenciando a relação com o contexto do falante.
Contextos e gêneros	Contextos costumam estar relacionados a diferentes <i>gêneros</i> discursivos. Os gêneros e contextos, assim como os eventos comunicativos e práticas sociais podem ser classificados, por exemplo, por esferas (público ou privado), modos (falado, escrito, multimedial, etc.), principais domínios sociais (política, mídia, educação, etc.), instituições ou organizações (Parlamento, Universidade, loja), papéis e relações dos participantes (médico-paciente, primeiro-ministro-membros do Parlamento), objetivos (transmitir ou buscar conhecimento, aconselhamento, atendimento, etc.) ou (inter)ações (tomada de decisão, governo, etc.), dentre outras. Os domínios podem ser agrupados em outros campos: a tomada de decisão, a ação e o controle coletivos (política, direito, administração, etc.), um campo simbólico referente à circulação de conhecimentos e crenças (mídia, educação, ciência, religião, etc.), um campo da produção (fabricação de bens) e um campo dos serviços (instituições de saúde, etc). Num nível mais baixo, os gêneros discursivos, eventos comunicativos ou situações sociais podem ser classificados por subtipos de ações: buscar conhecimentos a respeito de pessoas (entrevistas, interrogatórios), comunicar conhecimentos científicos (congressos, publicações em revistas especializadas, etc.), controlar as ações das pessoas (ordens, mandados de prisão, instruções normativas, manuais, etc.), etc.

Um modelo experiencial	Contextos são modelos mentais que representam situações comunicativas, e um tipo especial dos modelo mental que as pessoas constroem a partir das situações e entornos de suas vidas diárias. São esses os modelos de experiência. Esses modelos controlam a percepção da interação em progresso e possuem categorias básicas como: o ambiente espaço-temporal, os participantes e suas variadas identidades, os eventos ou ações em curso, bem como os propósitos daquele momento.
Controlador da produção e compreensão do discurso	Contextos, definidos como modelos mentais, controlam o processo de produção e compreensão do discurso e, portanto, as estruturas discursivas e as interpretações do discurso resultantes. Essa é a base cognitiva e a explicação daquilo que é chamado de influência da sociedade sobre o texto ou a fala, e o processo que garante que os usuários da língua consigam moldar seu discurso <i>apropriadamente</i> em relação às propriedades da situação comunicativa relevantes. Para uma análise, precisamos de uma teoria cognitiva articulada a partir de processos e representações. Curiosamente, esses processos estão inseridos em condições sociais e culturais mais amplas, compartilhadas por usuários da língua enquanto membros de grupos e comunidades.
Entidade dinâmica	Contextos, não são estáticos, e sim dinâmicos. Eles são construídos para cada situação comunicativa nova e, em seguida, <i>atualizados e adaptados</i> tendo em vista (a interpretação subjetiva das) restrições presentes da situação, incluindo o discurso e a interação imediatamente precedente. Em outras palavras, os contextos se desenvolvem em paralelo com a interação e (outros) pensamentos.
Modelos pragmaticamente e funcionais	A função fundamental do contexto é garantir que os participantes possam produzir textos ou falas <i>adequados</i> à situação. A língua é adaptada em sua interação discursiva aos ‘encontros’ socioculturais e cognitivos do momento. Uma teoria com essas características também torna explícitas as condições de felicidade usuais dos atos ilocucionários e as condições de adequação, a polidez e outras dimensões da interação (Austin, 1962; Searle, 1969). Portanto, uma teoria explícita do contexto também proporciona uma sólida base para várias abordagens em Pragmática.
Esquema mental	Contextos não representam situações sociais ou comunicativas completas, mas apenas - esquematicamente - aquelas propriedades que são continuamente relevantes. Uma teoria de modelo de contexto é ao mesmo tempo uma teoria da relevância do nível pessoal e interacional das interpretações da situação.
Recurso dêitico	Os contextos são <i>egocêntricos</i> . Isto é, são definidos por parâmetros que incluem um ambiente que é o espaço-temporal do ato de fala ou escrita em curso, do <i>Ego</i> enquanto falante ou ouvinte, de <i>outros</i> participantes, ou a quem ouve, bem como as ações sociais em curso em que propósitos específicos são força motriz. As propriedades desta natureza ‘egocêntrica’ dos contextos definem as condições das numerosas <i>expressões dêiticas</i> diferentes existentes em muitas línguas, tais como os pronomes pessoais, os demonstrativos, os tempos verbais, as preposições, as expressões de polidez e deferência e assim por diante. Tais parâmetros variam de língua para língua, assim como de falante para falante. Eles têm extensões metafóricas: estar ‘à esquerda de alguém’ pode ser uma definição espacial ou ideológica. O mesmo vale para a inclusão dos participantes em grupos contextuais e para a distinção de pertencer ou não pertencer a um grupo, expressada tipicamente pelos pronomes ideológicos nós <i>versus</i> eles.
Adequação	<p>Uma abordagem pragmática do contexto precisaria dar conta da relativa adequação relativa do discurso. Esse critério é paralelo ao da boa formação para a sintaxe, ao da significatividade para a semântica intensional e ao da verdade (satisfação, etc.) para a semântica extensional/referencial. A noção de adequação não é definida de maneira muito exata, e apenas conceitualiza que o discurso enquanto ação social é mais ou menos aceitável, correto, feliz, etc.</p> <p>O discurso e suas propriedades variáveis precisam coincidir com a definição normativa corrente da situação. Essa adequação pode ser definida para todos os níveis e dimensões de um texto ou fala, tais como a entonação, a seleção do léxico, a sintaxe, as expressões indiciais, os tópicos, os atos de fala, a distribuição de turnos, etc.</p>

Culturalmente variável	Os esquemas de contexto e suas categorias podem variar culturalmente e, assim, definir condições de adequação diferentes para o discurso em sociedades diferentes. Embora algumas categorias contextuais possam (ou precisem) ser universais, como é o caso do Falante e de vários tipos de Destinatários, bem como do Conhecimento. Já outras podem variar culturalmente, por exemplo, as propriedades sociais específicas dos participantes. Nível social, poder e parentesco são propriedades relevantes aos Participantes nos esquemas de contextos de muitas culturas - controlando, por exemplo, muitas expressões de polidez e deferência - outras categorias contextuais podem ser mais circunscritas, e outras ainda são provavelmente irrelevantes em qualquer lugar.
-------------------------------	---

Fonte: Adaptado de Van Dijk (2012, p.34-44)

Cada uma das visões de Van Dijk, umas mais do que outras, podem ser usadas como paradigmas investigativos das múltiplas concepções que o conceito de contexto possui para a investigação sociocognitivo-linguística das tarefas e grade de avaliação do exame oral do IELTS. Há também de se considerar a visão de Matthiessen e Halliday (2009), que apresenta relevantes pontos de vista sobre a natureza da definição de contexto que complementam os pontos de vista de Van Dijk (2012) já listados, porque fornecem insumos para construção de modelo mental, tão defendido pelo linguista holandês.

Halliday (1973), fundador da linha de pesquisa linguística sistêmico-funcional, pesquisa as funções da linguagem e seus sistemas desde seus trabalhos inaugurais. Especificamente, ao falar sobre contexto, ele, junto de Matthiessen, o distinguem entre situacional e cultural.

Contexto é um sistema semiótico de alto nível [sociocognitivo] no qual a linguagem está “encaixada”. Mais especificamente, a linguagem está encaixada em um contexto de cultura ou sistema social. Qualquer instância de linguagem como texto é encaixada no seu próprio contexto situacional. Contexto é uma matriz ecológica para ambos o sistema geral de língua e para textos específicos. Ele é realizado por meio da língua; e sendo realizado pela língua, por meio da linguagem humana, significa que ele cria e é criado pela linguagem. (MATTHIESEN; HALLIDAY, 2009, p.88)

A natureza interdependente da criação de linguagem por meio de contexto e viceversa é um dos pontos de alinhamento entre Matthiessen e Halliday (2009) e Van Dijk (2012), uma vez que ambos apontam para a necessidade de conceber práticas sociais como simbioticamente ligadas à função da linguagem, uso de lexemas, estruturas sintáticas características, além de comportamentos verbais e não-verbais interacionalmente situados. Ramos (2007) também aborda contexto, evidenciando que os recursos cognitivos suscitados

para mobilizar um determinado construto em uma prova estão intimamente ligados a ambientes reais, culturas específicas, textos específicos e gêneros discursivos específicos.

Ainda em Ramos (2007), há menção a Meurer (2006), cujas definições de contexto são duas, assim como as de Matthiessen e Halliday (2009). Alinhada à Linguística Sistêmico-Funcional, e aprofundando o texto dos primeiros funcionalistas, o primeiro conceito é o situacional, que “refere-se aos elementos mais imediatos que têm impacto sobre o texto e é especificado em termos de três componentes ou variáveis contextuais: o campo, as relações e o modo.” (MEURER, 2006, p.168-9 apud RAMOS, 2007, p.76). O segundo conceito é o cultural, que, para Meurer (2006) representa “os elementos de ordem social mais ampla - para além do contexto da situação - que de uma forma ou de outra influenciam os textos e são por eles influenciados.” (MEURER, 2006)

Com base nos postulados anteriores, podemos dizer que notáveis questões dignas de investigação vêm à tona. Por exemplo, de acordo com Van Dijk (2012), o contexto é inerente à natureza social e rege não só o léxico, mas o caráter geral das situações comunicativas em que nos inserimos. Meurer (2006) aponta que as relações entre os interactantes evidenciam muito a natureza do discurso co-construído, bem como a cultura que os rodeia. Como ignorar, ou lidar com tais afirmações, por exemplo, num teste que afere, por meio da mesma prova oral, dois propósitos sociais diferentes e, portanto, pressupor contextos também diferentes?

O parágrafo anterior traz apenas um dos muitos exemplos a serem desenvolvidos quanto ao alinhamento existente entre a TSC, as definições de contexto de Van Dijk (2012), Matthiessen e Halliday (2009), Ramos (2007) e Meurer (2006). A perspectiva sociocognitiva avaliativa e a perspectiva sociocognitivas linguística complementam-se não só em termos de diretrizes analíticas, isto é, o que elas se propõem a investigar, mas também em termos teórico-ideológicos. Por isso, utilizaremos os conceitos citados neste capítulo para complementar a análise da prova que será feita por meio do modelo sociocognitivo de Weir (2005), a ser apresentado no capítulo subsequente.

Capítulo 3 - Sociocognição e exames - Um modelo Sociocognitivo para investigações de Validade

Este capítulo visa apresentar o modelo sociocognitivo de validação de testes de Weir (2005), uma de nossas fontes de consulta e aporte teórico para a análise sociocognitivo-linguístico-avaliativa da validade de contexto das tarefas e grade de avaliação do exame oral do IELTS. Neste capítulo, abordaremos conceitos basilares para a investigação, como construto e validade, que compõem parte do modelo de validade de contexto na visão de Weir (2005). Apontaremos, também, como concebemos contexto de um posto de vista avaliativo, o que significa operacionalizar um construto, uma vez que esse é o objetivo das tarefas do exame, e validade, que, em poucas palavras, é a característica que mensura o alinhamento do exame a seus propósitos. Discorreremos sobre a importância do estudo de validade para exames e apresentaremos, por fim, o diagrama sociocognitivo de Weir (2005), cujos aspectos estão especificados na seção subsequente à apresentação do modelo.

Após especificação do modelo, entraremos em detalhes sobre o que dele usamos em nosso trabalho e definimos cada um dos itens presentes no modelo sociocognitivo de validade de contexto.

3.1 O que é construto em exames de proficiência?

Bachman (1990), Brown (2000) e Weir (2005) apontam que uma das maiores dificuldades de um exame é referente ao que ele deseja aferir, isto é, delimitar o que deve compor o seu construto e operacionalizar o construto-alvo da melhor maneira possível.

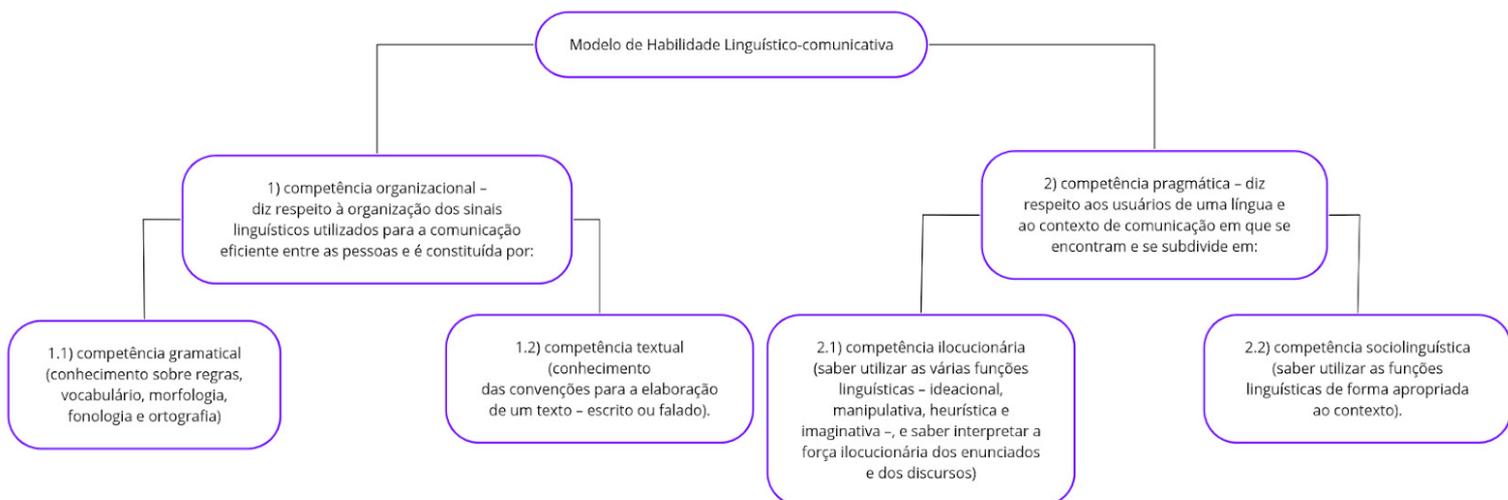
Bachman (1990) defende que simular um contexto que assegure uma precisa operacionalização de construto é um dos passos para um exame confiável. Para que os resultados de um exame sejam confiáveis, eles devem ser válidos, e para serem válidos, o construto, isto é, uma habilidade a ser testada, deve ser definida a partir de uma teoria geral de proficiência ou de uma ementa de habilidades, ou conteúdo, por exemplo. O construto, ou habilidade, são assuntos notavelmente problemáticos, aos olhos desse autor, por se tratarem de conceitos cognitivos que, quando operacionalizados, muito provavelmente não trarão à

tona o necessário para uma avaliação justa do candidato e de sua real capacidade. Para operacionalizar um construto por meio de um teste, Bachman (1990) propõe três passos:

- (1) identificação e definição do construto teórico;
- (2) definição do construto de maneira operacional;
- (3) estabelecimento de procedimentos para quantificar as observações do exame.

Em casos de construtos gerais, ou seja, habilidades voltadas à comunicação e interação em situações não específicas com a língua, um construto geral deve ser definido para fins avaliativos. Um dos construtos gerais mais popularmente usados é o de Bachman (1990), cuja referência é diretamente feita por Galaczi e ffrench (2011) quanto à concepção de língua e construto que os exames de proficiência C2 Proficiency e C1 Advanced se propõem a avaliar. O modelo, tal como Bachman (1990) coloca, não deve ser visto como estático, e sim dinâmico. Isto é, as alocações das habilidades nele presentes são meramente ilustrativas, podendo se relacionar de diferentes maneiras.

Figura 1: Modelo de habilidade Linguístico-comunicativa



Fonte:(BACHMAN, 1990); (QUEVEDO-CAMARGO, 2019, p. 32-33)

Outra definição de construto é a de Brown (2000), que define construto como uma característica, habilidade, demonstração de proficiência, ou mesmo potencial capacidade cognitiva, alocada no cérebro humano, e definida a partir de conceitualizações teóricas. As conceitualizações teóricas mencionadas por Brown (2000), tal como em Bachman (1990),

referem-se às reflexões teórico-metodológicas da natureza cognitiva do construto e dos instrumentos e ferramentas que pesquisadores podem utilizar para diferenciar um construto de outro que possa ser similar.

A perspectiva de Brown nos fornece uma visão mais holística de construto, mas que, ainda assim, segundo o autor, é problemática devido à sua inerente natureza psicológica. Por ser fisicamente intangível, mas empiricamente aferível, encontramos, ao fim da observação, com uma materialização do construto em potencial que pode estar incompleta. Brown (2000), assim como Bachman (1990), concebe que a característica cognitiva a ser operacionalizada, não demonstra toda a extensão do potencial pragmático de um construto cognitivo. Ambos os autores concordam que, para uma aferição alinhada com a operacionalização prática de um construto, as tarefas do exame devem ser calcadas em operações relevantes para que o construto emerja.

Weir (2005) define construto a partir do conceito de traço. O traço é definido como a manifestação verbal do construto psicológico subjacente à performance do que desejamos aferir. Para tal aferição é necessário um ou mais método(s) que são o(s) meio(s) operacional(is) que podemos utilizar para avaliar a performance dos candidatos.

Os três autores têm como ponto pacífico que, ainda que organizemos todos os instrumentos necessários para a aferição de um determinado construto, devemos sempre nos perguntar se o teste, ou exame, em questão, realmente está aferindo o que ele se propõe a aferir. Em outras palavras, devemos nos perguntar se há relação palpável entre um construto e os instrumentos utilizados para operacionalizá-lo, avaliá-lo, assim como sua relação com seus contextos sociais.

Em nossa pesquisa, cujo foco é a validade de contexto do exame oral do IELTS, a ser definida adiante, focamos na relação existente entre os instrumentos utilizados para provocar o uso do construto-alvo e os contextos que permeiam seu uso cotidiano.

3.2 O que é validade?

É de se imaginar que, ao investigar a validade de contexto do exame oral do IELTS, devemos definir o que compreendemos pelo conceito de validade. Discussões sobre validade antecedem os autores referenciados anteriormente para definir o termo construto. Lado (1961), por exemplo, é um dos primeiros autores a discutir validade em contexto de aquisição de segunda língua. Weir (2003; 2005), à sua maneira, é pioneiro ao cunhar uma teoria sociocognitiva avaliativa em que a validade não está somente ligada à precisão com que exames conseguem mobilizar e aferir um construto, mas também em como eles conseguem fazer com que este mesmo construto desvele-se durante a avaliação, e como os dados resultantes são capazes de refletir tal desvelamento em termos práticos. Isto é, como os dados obtidos podem ser utilizados por instituições, empregadores, universidades, dentre outros.

Assim sendo, tal como Bachman (1990), cremos que a validade de um teste é sua característica mais importante e também aquela que mais interfere no quão alinhado aos seus propósitos, o quão significativo e útil um teste pode ser. No entanto, diferentemente de Bachman que afirma que validade é uma qualidade concernente “à interpretação e uso de um exame”³⁰ (BACHMAN, 1990, p. 25), pensamos, tal como Weir (2005), que validade é uma característica indissociável das demais validades presentes em um teste. Isto não é dizer que Bachman (1990) não concebe que validade e confiabilidade, por exemplo, estão intimamente relacionadas, e que ao sacrificar uma em detrimento da outra, por vezes, perdemos ambas; contudo, o autor não calca seu modelo de investigação de exames de línguas no conceito de validade em si.

Nesse sentido, optamos por partir da pesquisa de Weir (2005) em que, tal como em Messick (1995), o conceito de validade é central e base para a investigação de exames. Dessa forma, a partir de vários tipos de validade, o autor nos guia para aspectos outros da prova. No modelo sociocognitivo de Weir (2005) para a investigação de exames orais, a base teórica de nossa pesquisa, há cinco tipos de validade, as quais listamos por meio de breves definições a seguir.

³⁰ no original “its interpretation and use”

Quadro 2 - Resumo dos tipos de Validade do modelo sociocognitivo de Weir (2005) para exames Orais

Tipos de validade presentes no modelo sociocognitivo de Weir	Definição
Validade de contexto	<p>A validade de contexto relaciona-se à maneira com que as tarefas de um teste assemelham-se ao uso de uma habilidade ou construto.</p> <p>Se as tarefas de um teste refletirem as situações que as operacionalizam na vida real, acabará por facilitar a aferição do que um candidato pode atingir por meio da língua-alvo. Para tal, é necessário entender e incorporar as condições normais de funcionamento de um determinado ambiente para que se possa operacionalizar seu respectivo construto. Do contrário, seria imprudente fazer julgamentos sobre o uso da língua-alvo de um determinado candidato da parte do examinador. (WEIR, 1993; WEIR, 2005)</p>
Validade com base em teoria	<p>Desenvolvedores de testes precisam estar cientes das principais teorias concernentes ao processamento de linguagem que subjazem às várias operações necessárias para o uso da língua na vida real. A operacionalização de um construto em um teste deve embasar-se e replicar os processos cognitivos subjacentes ao uso da língua. (WEIR, 2005)</p>
Validade de pontuação	<p>Para Weir (2005), validade de pontuação é uma espécie de hiperônimo para tratar de tudo aquilo que concerne à confiabilidade e validade de um teste. Assim sendo, a validade de pontuação trata dos aspectos referentes à pontuação de um teste e sua consistência em termos de amostra de conteúdo e vieses avaliativos. Em outras palavras, este tipo de validade ocupa-se de tentar garantir uma pontuação livre de falhas e, conseqüentemente, a confiabilidade do exame para decisões que possam depender dele. (WEIR; 2005)</p>
Validade consequencial	<p>A validade consequencial relaciona-se com as conseqüências que as interpretações e usos de um teste podem acarretar na vida real. Ela divide-se em três subtipos:</p> <p>[1] Validade diferencial</p> <p>É concernente à maneira como o teste favorece ou discrimina o tratamento dos candidatos, viés(es) ideológico(s) e/ou construtos (im)pertinentes a um teste.</p> <p>[2] Efeito retroativo</p> <p>Lida com a maneira como testes afetam o ensino.</p> <p>[3] Efeitos na sociedade</p> <p>O efeito de um teste na sociedade é um dos aspectos mais difíceis de serem investigados. Isso se dá porque é necessário empenho investigativo que vá além do teste e seus usuários. (WEIR, 2005)</p>

Validade relacionada ao critério	<p>A validade relacionada ao critério divide-se em dois tipos.</p> <p>A validade concorrente procura por um critério que sirva de indicador da habilidade testada. A pontuação do teste pode estar relacionada a outro tipo de mensuração, normalmente mais antiga, longa, e estabelecida, o que pode prejudicar uma avaliação precisa. (WEIR; 2005)</p> <p>A validade preditiva relaciona-se ao poder preditivo de um instrumento em relação à maneira como ele operacionaliza seu construto e realiza sua mensuração (WEIR; 2005).</p>
----------------------------------	--

Fonte: (WEIR, 2005)

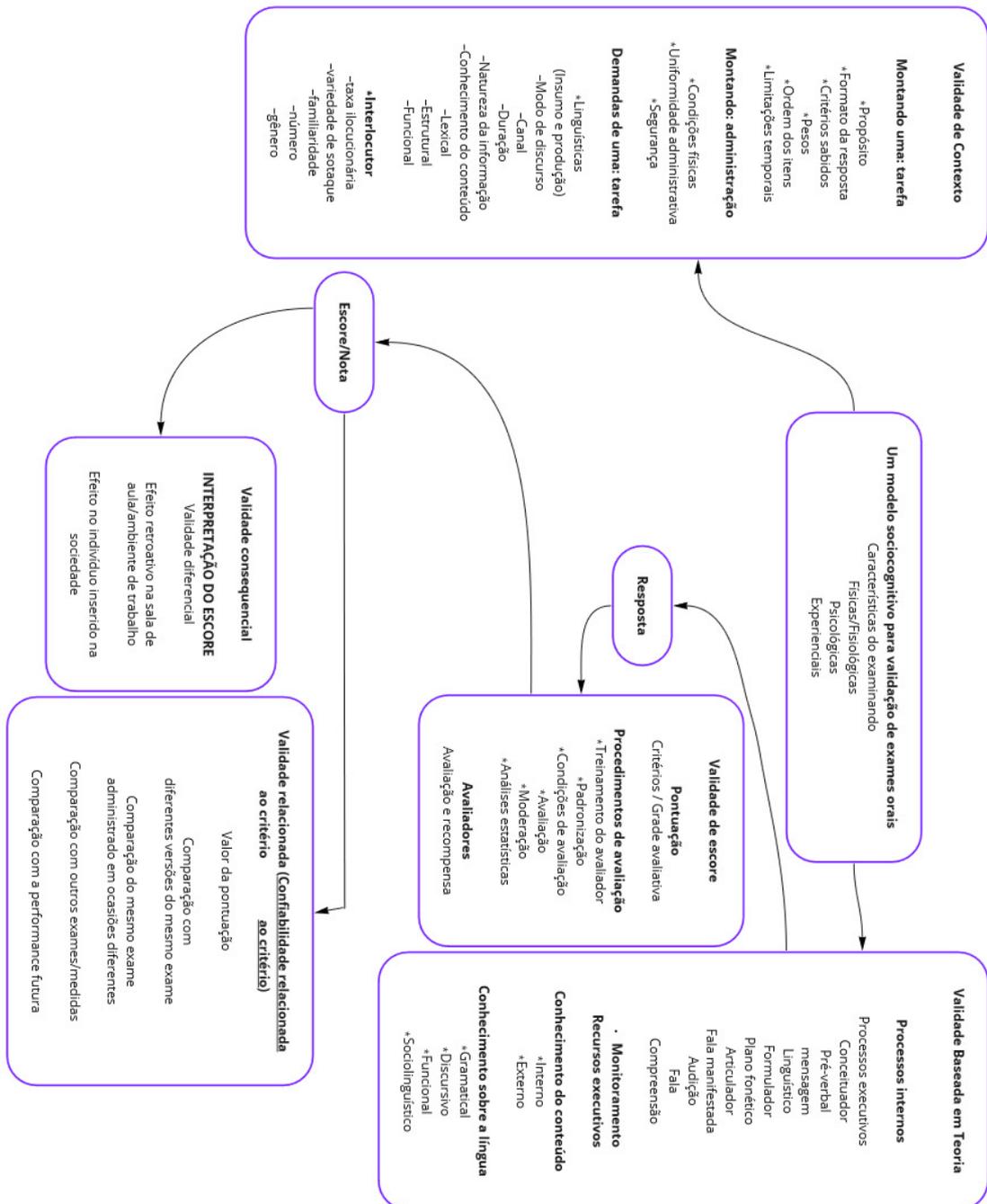
Os cinco tipos de validade anteriormente descritos servem de base para o modelo sociocognitivo de Weir (2005) transposto na Figura 2.

Do modelo em questão, há partes que não temos tempo, nem recursos, para investigar. Não somos capazes, por exemplo, de investigar todos os tipos de validade nele presentes pelo exíguo tempo característico do mestrado. Por essa razão, optamos por ater-nos à validade de contexto, somente, ainda que com ressalvas. A primeira ressalva é relativa ao insumo e à produção da tarefa (*input* e *output*) em decorrência dos entraves causados pela pandemia de COVID-19, o que nos impossibilitou, e ainda impossibilita, de conduzir uma pesquisa de campo com candidatos da prova para conseguir amostras de produção (*output*). A segunda ressalva é referente à administração da prova, a qual só têm acesso os envolvidos na aplicação do exame, são eles: as escolas e/ou instituições que recebem os candidatos e o BC, incluindo seus funcionários e afiliados. A terceira ressalva é referente ao mediador, pois exigiria de nós, assim como na primeira ressalva, uma série de entrevistas com diversos interlocutores que aplicam a prova. Feitas as ressalvas, apresentamos o que nos cabe do modelo sociocognitivo de validade de contexto.

Propomo-nos a investigar se e como os elementos sociocognitivos relacionados à validade de contexto do exame oral do IELTS estão alinhados à perspectiva sociocognitiva de validade de contexto com base nos propósitos acadêmicos e gerais, tendo em vista a investigação de contexto sociocognitiva a ser calcada no referencial do capítulo anterior. Para tal, adaptamos o modelo original de validade de contexto de Weir (2005) a um modelo adequado ao tipo de investigação que utilizamos em nossa pesquisa, uma vez que documentos e artigos serão nossas primordiais fontes de coleta de dados. Nas figuras 3 e 4, apresentadas

na seção a seguir, temos as diferenças entre os dois modelos, o original, e o nosso, adaptado.

Figura 2 - Modelo Sociocognitivo



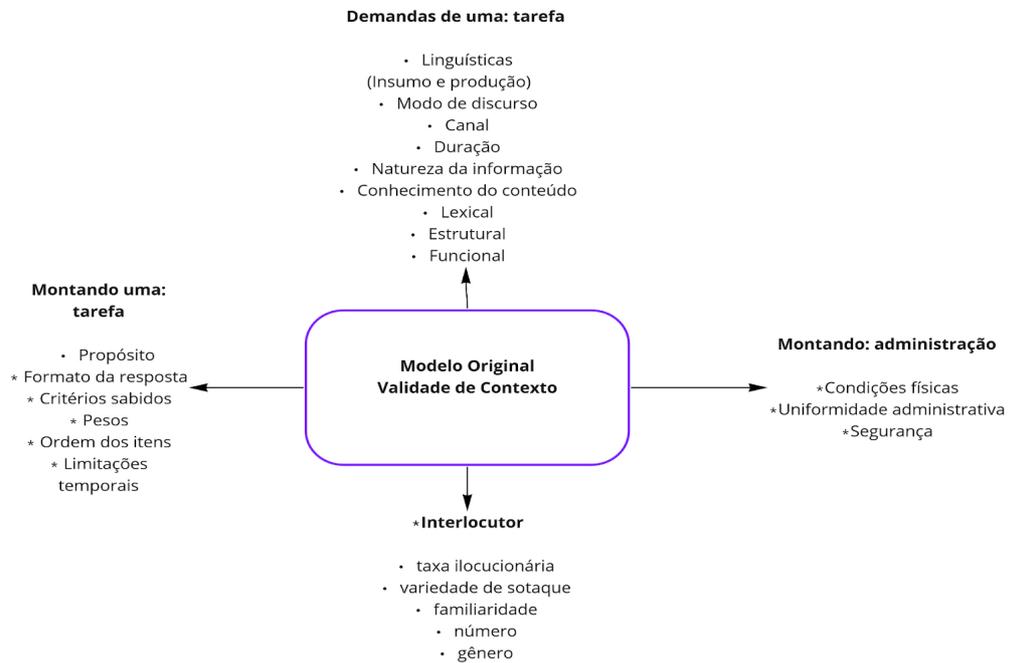
Fonte: (WEIR; 2005, p.46)

3.3 Definindo os critérios de análise do modelo sociocognitivo de Weir (2005)

Iniciamos esta seção retomando as definições de construto, que é uma habilidade ou competência psicológica que se deseja avaliar, e validade de contexto, que é a relação existente entre a operacionalização do construto-alvo da vida real em relação à sua contraparte operacionalizada em um teste.

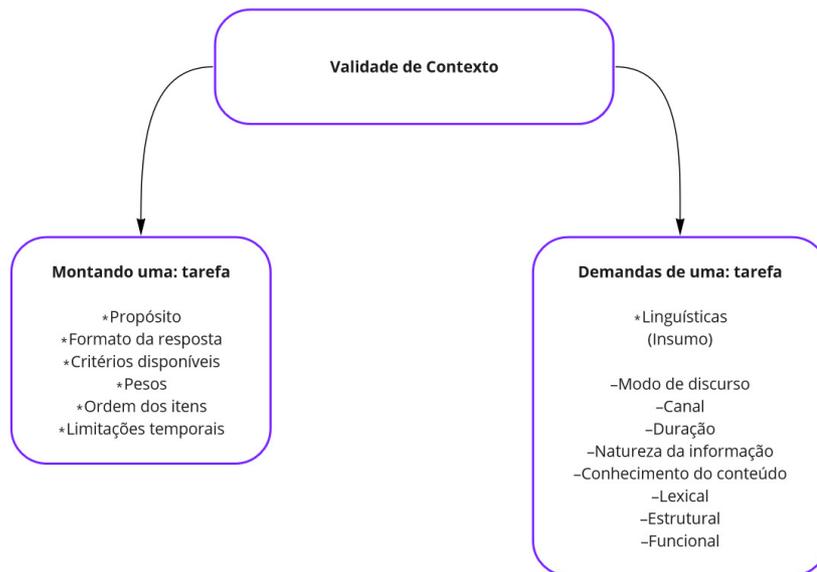
Para que o construto seja válido, ele deve estar adaptado ao contexto cujo exame se propõe a operacionalizar e avaliar. Weir (1993; 2005) ressalta que caso o cenário proposto pela tarefa do exame não reflita, dentro das limitações inerentes a um exame, seu contexto real da maneira mais verossímil possível, há de se reconsiderar a estrutura da tarefa e (re)investigá-la a fim de aferir o alinhamento entre o propósito do teste e as técnicas de operacionalização, mensuração e avaliação do construto. Em outras palavras, o construto a ser investigado por meio de um exame deve ser operacionalizado com base no propósito e contexto em que se deseja inserir o candidato (BROWN, 2000).

Figura 3 - Modelo Sociocognitivo Original de investigação de Validade de Contexto



Fonte: (WEIR, 2005, p. 46)

Figura 4 - Modelo Sociocognitivo Adaptado de investigação de Validade de Contexto



Fonte: (WEIR, 2005, p. 46)

Creemos que nossa costura teórica subsidia nossa visão focal uma vez que a validade de contexto procura investigar como os parâmetros contextuais presentes num exame de proficiência relaciona-se ao propósito, tarefas, níveis de proficiência e grade de avaliação.

A seguir, apresentamos um resumo das definições de Weir (2005) para cada um dos tópicos elencados em nosso modelo. Começando pelos comandos, parte integrante do propósito de uma tarefa.

3.4 Instruções, enunciados e comandos

Partindo de nossa adaptação do modelo de validade de contexto, primeiramente, ao investigar tarefas, devemos procurar compreender a maneira como ela foi organizada e se tal organização evidencia o propósito da tarefa. Os pontos tidos como mais relevantes à tarefa e ao candidato são elencados por Khalifa (2003):

- A instrução está clara sobre o que os alunos devem fazer?
- A instrução foi escrita com as frases mais breves e simples possíveis?
- A instrução está gramaticalmente correta?
- A instrução foi escrita corretamente?
- A instrução está no primeiro idioma ou no idioma-alvo?
- A instrução é clara sobre a quantidade de tempo gasto em cada parte da tarefa?
- A tarefa exigirá diferentes tipos de resposta? Nesse caso, pode ser necessário fornecer instruções específicas separadas para cada tipo de resposta necessária.³¹

(KHALIFA, 2003, p.73 apud WEIR, 2005, p. 57)

Nota-se que, no original, o termo utilizado nos excertos é *rubric*, que segundo o dicionário online da Universidade de Cambridge (2021) resume-se a um conjunto de instruções ou explicações concernentes, principalmente, a um exame ou livro. Devemos, porém, desambiguar um conceito que soa como um hiperônimo, pois possíveis problemas afloram a partir das diferenças entre *instrução*, *enunciado* e *comando*. Tratam-se de três categorias distintas: 1) Instrução “é uma orientação que indica para o aluno de que *forma* ele

³¹ No original: • Is the rubric clear on what students have to do?; • Is the rubric written in as short as possible, simple sentences? • Is the rubric grammatically correct? • Is the rubric spelled correctly? • Is the rubric in the First Language (L1) or the Target Language (TL)? • Is the rubric familiar to the students? • Is the rubric clear about the amount of time to spend on each part of the task? • Will the task require different types of response? If so, it may be necessary to provide separate specific instructions for each type of response required.

deve responder o item (questão).” (FERNANDES, 2015, p.5) - um exemplo de instrução seria: “Após ler e analisar o poema, assinale a alternativa que completa corretamente a afirmação.” (FERNANDES, 2015, p.5); 2) já o enunciado “traz em si o estímulo que provoca a resposta. É uma situação problema expressa como afirmativa ou pergunta e explicita claramente a base da resposta – o que se exige do avaliando – e como ele deve proceder – o comando da resposta.” (GOVERNO DE MINAS GERAIS, 2021, p.8) - exemplo de enunciado seguido de comando: “A argumentação do eu poético nesse soneto de Gregório de Matos faz com que, pela lógica, ele se apresente perante Deus como alguém...” (FERNANDES, 2015, p.4). Nesse último exemplo, o enunciado retoma o texto motivador utilizado na tarefa e, a partir das instruções iniciais, elucida o comando, que é identificar como o eu lírico do soneto se apresenta perante Deus.

Em resumo: *[a](s) [instruções, enunciados e] comando(s) do teste são preciso(s) e acessível(is)?* (WEIR, 2005)

3.5 Propósito

Aos candidatos de um teste deve ser dada uma ideia clara e inconfundível sobre o que o comando da tarefa espera deles para que possam escolher as estratégias mais adequadas para construir uma resposta satisfatória a partir do insumo disponibilizado. Em um teste oral, o propósito ajuda os falantes a definir a estrutura e foco da interação. É improvável, no entanto, que um teste seja capaz de operacionalizar a variedade de condições e operações cognitivas necessárias presentes em uma situação da vida real.

Ademais, o propósito pelo qual os candidatos estão realizando uma determinada tarefa deve sempre estar claro para o candidato (além da demonstração de habilidades por meio de várias características da fala).

Em resumo: *o objetivo do teste é inequivocamente claro para o candidato? É um propósito apropriado?* (WEIR, 2005)

Logicamente, a instrução, o enunciado, o comando e o propósito de uma tarefa são arranjos para facilitar o formato da resposta esperado, que é o próximo item de nossa lista.

3.6 Formato da Resposta

O formato de resposta é, em Weir (2005), o método que utilizaremos para gerar os dados referentes ao construto. No entanto, há de se considerar que o método usado para testar a habilidade do idioma pode afetar a pontuação do aluno (ALDERSON et al, 1995)

Faz-se necessário investigar o formato da resposta para que possamos investigar o formato e/ou método de avaliação, e se esse está alinhado ao propósito de avaliação de um exame (WEIR, 2005). A escolha feita quanto ao formato da resposta pode afetar significativamente o processamento cognitivo da parte do candidato. Nesse sentido, o método de geração de dados não deve intervir adversamente no processamento cognitivo do construto-alvo. Para tal, devemos considerar a maneira como processamos as interações em seus contextos originais e, à medida que as interações simuladas no exame se distanciarem do ambiente evocado, mais desalinhadas elas se tornam, e conseqüentemente, perdem no quesito validade de contexto.

O formato da resposta também está relacionado ao tipo de habilidade que o exame deseja aferir. Luoma (2004) defende que na testagem “direta”, lidamos com a habilidade em questão diretamente. Num exame oral cuja entrevista é direta, o mediador-avaliador deve participar da interação face a face. De maneira geral, os resultados de testes diretos tendem a ser mais confiáveis que os de testes indiretos ou semi-diretos.

Em exames semi-diretos ocorrem por meio de diferentes insumos cujas mídias podem variar. O candidato deve reagir ao insumo disponibilizado e sua resposta será gravada em um dispositivo eletrônico (ex: computador ou gravador). Um teste indireto é aquele que procura testar, tal como descrito na introdução, pronúncia através da escrita.

Em resumo: *existe alguma evidência de que o formato de resposta do teste pode afetar o desempenho de seu candidato?* (WEIR, 2005)

3.7 Critério disponíveis

Os critérios, em poucas palavras, são os tópicos elencados na grade de avaliação para guiar o examinador no processo de mensuração do desempenho dos candidatos de um exame. Os candidatos devem saber como serão julgados, isto é, o que é necessário para responder às questões de compreensão. Tais informações são cruciais para a realização de um exame e devem estar disponíveis para os candidatos e professores antes de sua realização. Isso inclui todas as informações já publicadas, como: a maneira com que as tarefas são pontuadas, critérios de correção, etapas usadas para pontuação e como as pontuações dos itens são combinadas na pontuação do teste, devem estar prontamente disponíveis (WEIR, 2005).

Se algum dos critérios não for usado na marcação, afetam-se os mecanismos de planejamento e execução do processamento cognitivo envolvido na conclusão da tarefa.

Em resumo: *os critérios a serem usados para pontuar as tarefas da prova são explícitos para os candidatos e os examinadores?* (WEIR, 2005)

3.8 Peso

O estudo do peso de um teste se ocupa de investigar as atribuições numéricas referentes à mensuração e à distribuição de diferentes pontos mínimos e/ou máximos a um item, tarefa, ou demais componentes de um teste para alterar a relação contributiva entre uma tarefa e outras dentro de um mesmo teste. Se diferentes partes do teste forem pontuadas de formas diferentes, então, o tempo e/ou notas a serem atribuídas também deverão refletir essa diferença (WEIR, 2005). Tais diferenças devem ser evidentes para os participantes do teste para que eles possam alocar seu tempo de acordo durante a fase de processamento de metas e arranjos da tarefa.

Em resumo: *de que maneira um item que testa um aspecto X é mais ou menos importante do que um item que testa um aspecto Y? Há alguma ponderação para os diferentes componentes de um teste que tenha sido devidamente justificada?* (WEIR, 2005)

3.9 Ordem dos itens

A ordem dos itens deve refletir a maneira como essas habilidades e estratégias são mobilizadas no processamento cotidiano quanto à leitura ou ao construto-alvo. Em um teste de fala ou escrita, pode-se organizar a ordem em que as tarefas ocorrem por razões lógicas ou afetivas (BUCK, 2001).

Em resumo: *os itens e as tarefas de um teste estão em uma ordem justificável?* (WEIR, 2005)

3.10 Limitações de tempo

Ao desenvolver um teste, seus elaboradores devem sequenciar os textos (quando houver), enunciados, comandos e tarefas a fim de garantir que haja tempo suficiente para todas as atividades; se a alocação do tempo não for devidamente planejada, o desempenho do candidato pode tornar-se imprevisível. Por exemplo, se muito tempo é dado em um teste de leitura, ou não é estritamente controlado por seção, os candidatos podem simplesmente ler uma passagem e suas respectivas perguntas, anteriormente destinadas a testar a capacidade de processar o texto rapidamente (ou seja, seletivamente e rapidamente) e deixar de ativar o construto-alvo daquele contexto de leitura (WEIR et al, 2000). Isso resulta em um processamento cumulativo, linear e cuidadoso, em vez de um processamento seletivo rápido. As decisões relacionadas ao tempo têm um impacto claro no processamento e, portanto, na validade baseada na teoria de nossas tarefas de teste. (WEIR, 2005)

As tarefas podem ser respondidas de forma satisfatória no tempo permitido? Se houver pouco tempo disponível, haverá estresse e não estaremos obtendo o melhor desempenho do aluno (WEIR, 2005).

Não obstante, os candidatos devem saber quanto tempo eles têm *em cada parte de um teste*. A divisão do tempo também deve refletir a importância que cada parte do domínio da língua que está sendo testado tem para seus avaliadores. No passado, pensava-se que as tarefas de teste com tempo limitado seriam apenas uma representação limitada do que os escritores costumam fazer ao criar um texto escrito, por exemplo. A performance dos candidatos não era interpretada como uma produção representativa de sua melhor performance (WEIR, 2005).

Há experimentos que apontam para o contrário. Kroll (1990) comparou ensaios escritos em sala de aula com tempo limite e ensaios escritos em casa variando de 10 até 14 dias. De maneira geral, com base nos resultados, o tempo não colabora muito para que os alunos tenham melhor desempenho. No experimento de Kroll, erros de sintaxe permaneciam notavelmente semelhantes em ambos - assim, como em suas habilidades organizacionais.

Nos testes orais, no entanto, diferentes questões relacionadas ao tempo se apresentam. A fala cotidiana depende de tempo para ocorrer. Pausas e hesitações possuem limites culturais de aceitabilidade. Apenas uma quantidade relativamente limitada de silêncio é tolerada em algumas culturas, enquanto em outras, esse limite pode ser maior. Falantes com boa pronúncia e fluência também são, normalmente, tido como mais hábeis que aqueles que apresentam mais pausas e dificuldade de pronúncia (WEIR, 2005).

Dar destaque ao processamento do tempo inerentemente ligado à cultura da língua-alvo em testes de idiomas pode encorajar mais práticas que abordem tal questão na sala de aula de idiomas. Outro aspecto a se considerar é o tipo de preparação para uma prova oral; por exemplo, se o candidato tiver mais tempo para preparar seu discurso detalhadamente em vez de ter que falar espontaneamente, pode-se ter gritantes diferenças de desempenho (GOFFMAN, 1986; GUMPERZ, 1980; WEIR, 2005).

Weir (2005) também aponta que devemos nos atentar para a diferença de duração de turnos. Turnos curtos são mais comuns no cotidiano; geralmente, frases mais espontâneas com menor rigor gramático-normativo permeiam a fala. Turnos longos, por exemplo, como apresentações orais ou palestras, por definição, requerem mais planejamento e, portanto, tendem a ser mais bem preparadas, com frases mais complexas e maior rigor gramático-normativo.

Em testes em que turnos longos são uma característica, deve-se decidir o tempo de planejamento que deve ser concedido aos candidatos. O tempo deve ser proporcional à dificuldade da tarefa, do contrário, pode prejudicar o processo de formulação da resposta do candidato. Há de se considerar até que ponto um candidato deve interagir com o insumo, on-line ou em tempo real. Para que o nível de resposta seja considerado como central na dificuldade de uma determinada tarefa, o desempenho bem-sucedido da tarefa deve exigir que o candidato processe as informações essenciais da tarefa de uma maneira relativamente imediata, tal qual fora da situação simulada (WEIR, 2005).

Em resumo: *o tempo para cada parte do teste, por exemplo, preparação e conclusão,*

é apropriado? (WEIR, 2005)

3.11 Demandas Linguísticas de uma tarefa

3.11.1 Modo de discurso

Dentro do discurso oral há uma grande variedade de interações passíveis de simulação num exame oral, impossibilitando-nos de utilizar um termo geral que cubra todas as suas diferenças. Dessa forma, utilizamos os modelos mais comuns em provas de proficiência cujo exame oral está presente, segundo Weir (2005). São eles, primeiramente, a conversa e a entrevista.

A primeira diferenciação que o autor aborda quanto ao discurso falado é concernente às diferenças entre a entrevista formal com um aluno e o examinador e uma conversa entre colegas (VAN LIER, 1989). O alcance funcional obtido por cada uma dessas pode ser drasticamente diferente. Características como reciprocidade, assim como a responsabilidade pela manutenção da interação variam, principalmente em entrevista que os examinadores podem responder ao entrevistado de maneiras que não seriam consideradas aceitáveis em uma conversa natural - como, por exemplo, quando lhes é permitido desafiar o ponto de vista de um candidato. Para melhor exemplificar a discussão posta por Weir (2005), voltamo-nos para um referencial específico.

Discursos podem ser divididos segundo categorias contextuais, como a macrofuncional e a microfuncional. As macrofunções (COUNCIL OF EUROPE, 2001) são os traços funcionais marcantes de uma interação. As microfunções (COUNCIL OF EUROPE, 2001) são pertinentes a atividades interacionais subordinadas às rotinas interacionais. Essas funções discursivas evocam diferentes situações, como as descritas por Brown, Anderson, Shillcock e Yule (1984). Os autores, avaliam três situações: estáticas, dinâmicas e abstratas. Estáticas são relativas às descrições, dinâmicas à troca de informações e abstratas às divagações.

O'Sullivan (2002) é outro dos autores que elenca categorias interacionais que auxiliam a investigação de macro e microfunções de uma tarefa. A primeira é a categoria informacional, que atribui o tipo de informação e função a serem produzidas pelo candidato, que pode ser concreta ou abstrata. Ela é seguida da interacional, que atribui funções específicas aos interactantes e seus contextos suscitados por meio das tarefas e formato do exame. A terceira trata do gerenciamento das funções interacionais que investiga como e quais interações foram trazidas à tona.

Diferentes modos discursivos e funções são evocados com base nas definições anteriores. Brown e Yule (1983) e Bygate (1987) são exemplos de pesquisadores que organizam funções por complexidade. Bygate (1987), por meio de sua definição de *rotinas interacionais*, isto é, hábitos, ou *scripts* que criamos em nossa mente, organiza as interações de maneira temática e dificuldade cognitiva gradativa por meio de duas categorias: a primeira é a conversa-factual (*factually-oriented talk*), que compreende *descrição, narração, instrução e comparação*; a segunda é a conversa-avaliativa (*evaluative-talk*) que compreende *explicação, justificativa, produção de ideias e decisão*. Brown e Yule (1983) são os responsáveis pelas descrições da análise do comportamento discursivo co-construído dentro dos tipos de conversa definidos por Bygate (1987). Em ordem de complexidade, os autores organizam as funções *descrição, narração, instrução e comparação*, seguido de *explicação, justificativa, produção de ideias e decisão*.

Bereiter e Scardamalia (1987) complementam as funções elencadas por Brown e Yule (1983) com diferenciações entre as atividades mais complexas e simples: as primeiras categorias servem para revisar conhecimento; a segunda para transformar conhecimento. Vähäpääsi (1982) compartilha da primeira visão numa natureza informacional, categorizando-a como o ato de reproduzir informação. Já quando se trata de organizar informação, Vähäpääsi (1982) concebe que há maior demanda cognitiva porque as informações necessárias para mobilizar tais funções como conhecimentos hipotéticos, opinativos, ou especulativos, não encontram-se na esfera da reprodução de conhecimento.

A maneira como as tarefas mobilizam as funções citadas, ou demais funções contextualmente definidas ditam a validade de contexto das tarefas. Já a maneira como elas são avaliadas, dita o alinhamento entre a grade e as tarefas, juntamente do contexto.

Em resumo: *o modo de discurso é, afinal, apropriado para as habilidades ou estratégias que se deseja avaliar?* (WEIR, 2005)

3.11.2 Características do modo de discurso: reciprocidade na interação falada

O termo reciprocidade diz respeito à dimensão da interação interpessoal, isto é, a relação entre falante e ouvinte. Nesse caso, o foco está na troca de turnos, em quem tem o direito de falar, e como o compartilhamento da responsabilidade na manutenção de uma interação se dá. O grau de reciprocidade/participação em uma interação em andamento varia. As variações dependem da situação interacional que ocorre. Uma palestra, uma entrevista, e uma conversa apresentam papéis discursivos, turnos de fala e responsabilidade quanto à manutenção da interação completamente diferentes (GOFFMAN, 1986; WEIR, 2005).

Weir (2005) defende que em algumas situações, como em uma palestra formal, apenas o falante normalmente tem o direito de falar, sendo o próprio falante responsável por manter o discurso em andamento. Em uma conversa informal, tanto o falante quanto o ouvinte têm direito à fala e troca de turnos. Em uma situação de entrevista em um teste, o examinador domina a interação, pelo menos em termos de gerenciamento do que deve ser ministrado durante o exame, normalmente definidos pelas seguintes etapas: início da discussão, continuação/manutenção e conclusão.

Na maioria das situações de fala, a pessoa com quem estamos falando está na nossa frente e é capaz de nos corrigir se cometermos um erro. Ele ou ela, geralmente, pode mostrar concordância e compreensão - ou incompreensão e descontentamento. Os palestrantes devem prestar atenção em seus ouvintes e adaptar suas mensagens de acordo com suas reações. Se o ouvinte sinaliza que compreendeu, seria estranho persistir em uma explicação (WEIR, 2005).

Em resumo: *o modo de discurso é, afinal, apropriado para as habilidades ou estratégias que se deseja avaliar?* (WEIR, 2005)

3.12 Canal de comunicação

O canal de comunicação refere-se às vias comunicativas que regem o teste. Será um teste que se realizará face a face? Ou será um teste remoto, online? O canal de comunicação pode afetar significativamente o desempenho de um candidato em um teste.

As decisões referentes ao canal de comunicação têm que ser feitas sempre em relação à natureza e à quantidade de informação não-verbal que é desejável. Tratando-se de uma atividade de produção, por exemplo, podemos pensar em gráficos, tabelas, diagramas, etc. A escolha do leiaute, ou mesmo do tipo de letra, deve ser feita sob consideração cuidadosa dos desenvolvedores, pois isso pode afetar a percepção da tarefa por parte do candidato.

Em resumo: *o canal é apropriado para os requisitos de situação-alvo dos alunos que estão sendo testados?* (WEIR, 2005)

3.13 Duração

A duração conecta-se ao tempo de processamento de um teste, construto, ou habilidade. Assim como as limitações temporais afetam a tarefa, a duração também acarreta em consequências diretas à produção do candidato. Normalmente, fala-se de duração quando discutimos atividades que envolvem a compreensão auditiva. Contudo, olhar para a duração de uma atividade oral é crucial para o planejamento de um teste.

O comprimento do texto, segundo Weir (2005), tem um efeito potencial adicional em termos dos recursos cognitivos que podem ser acionados durante o processamento cognitivo da tarefa em questão. De maneira geral, quanto mais tempo os candidatos têm para lidar com o texto, há uma maior chance de mobilização cognitiva do conhecimento do idioma-alvo que pode ser necessário para processar a questão. Se os textos não estiverem tratando das demandas inerentemente ligadas aos recursos cognitivos que devem ser mobilizados, compromete-se a validade da questão.

Em resumo: *o comprimento do texto é apropriado para os requisitos da situação-alvo*

dos alunos que estão sendo testados? (WEIR, 2005)

3.14 Natureza da informação no texto

A natureza da informação de um texto refere-se ao seu tema. Em outras palavras, deve-se pensar se a informação contida no texto é abstrata (ética, amor, etc.) ou concreta (os objetos em uma sala, por exemplo) (COUNCIL OF EUROPE, 2001; WEIR, 2005) e o quão relevantes são para o alinhamento existente entre teste e propósito. A presença de informações abstratas não exclui a possibilidade de inserção de informações concretas, no entanto, há de se considerar que a informação abstrata pode ser cognitiva e linguisticamente mais complexa e mais difícil de processar.

Em resumo: *o tipo de informação é apropriado para os requisitos da situação-alvo dos alunos que estão sendo testados?* (WEIR, 2005)

3.15 Conhecimento de conteúdo necessário

O conhecimento do conteúdo necessário para completar uma tarefa específica afetará a maneira como ela é avaliada. O conhecimento necessário manifesta-se em duas instâncias: a primeira é a relação entre o conteúdo do texto e os conhecimentos prévios do candidato (conhecimento geral que pode ou não ser relevante para o conteúdo de um determinado texto que inclui conhecimento cultural), e a segunda é o conhecimento do assunto (conhecimento específico diretamente relevante para o tópico e conteúdo do texto) a ser considerado. Esse tipo de interação entre os recursos cognitivos do candidato e aqueles exigidos pela tarefa, que mais uma vez enfatiza a natureza simbiótica da validade baseada no contexto e na teoria.

Douglas (2000) defende que um teste para fins específicos é aquele em que o conteúdo e os métodos do teste são advindos de uma análise situacional do uso do idioma-alvo para um propósito específico, de modo que as tarefas e o conteúdo do teste sejam autenticamente representativos das tarefas na situação-alvo, permitindo uma interação entre a habilidade de linguagem do candidato e o propósito específico, que configura-se como o conhecimento do

conteúdo, de um lado, e as tarefas de teste, do outro. Tal teste nos permitiria fazer inferências sobre a capacidade do candidato de usar a linguagem no domínio de propósito específico (DOUGLAS, 2000).

As tarefas dos textos motivadores, se presentes, devem ser alinhados em termos de gênero, tarefa(s) retórica(s) e padrão(ões) de mobilização de processos cognitivos específicos e em um nível apropriado de especificidade. Isso se dá porque, para grupos heterogêneos de alunos, os elaboradores são forçados a selecionar textos mais gerais, que não é o caso de quando temos um grupo de controle mais homogêneo. Nos testes de Inglês para fins Acadêmicos (em inglês, EAP - English for Academic Purposes), isso nos leva a áreas como educação em saúde, habilidades e hábitos de estudo, ecologia, etc (WEIR, 2005).

Se uma tarefa for considerada pouco palpável, inadequada ou impossível, os candidatos não serão capazes de desempenhar o melhor de suas habilidades e podem contestar ou ignorar a tarefa.

Em resumo, segundo Weir (2005) devemos nos questionar as seguintes perguntas sobre o conhecimento do conteúdo:

- *O assunto está de acordo com a experiência e o nível provável da faixa etária dos alunos?*
- *O assunto é adequado para todos os alunos? O assunto não deve ser tendencioso de nenhuma forma (em termos de urbano/rural, menino/menina, cultural, etc.).*
- *Existem assuntos inadequados como guerra, morte, política e crenças religiosas que podem ofender ou angustiar alguns alunos? Isso deve ser evitado.*
- *Existe uma suposição muito grande de conhecimento cultural no assunto? Não deveria haver.*
- *O assunto do texto deve ser pouco perceptível socialmente, mas familiar, embora não muito familiar para os alunos, ou seja, o texto deve ser pouco perceptível socialmente, mas os alunos têm um esquema existente (estrutura mental organizada) para o assunto.*
- *O estilo de diálogo/passagem é familiar aos alunos? O tipo de texto deve ser semelhante ao do livro didático, por exemplo, descrição, narrativa, argumento.*
- *O conteúdo do assunto é apropriado para os requisitos da situação-alvo dos alunos que estão sendo testados?*³² (WEIR, 2005, p.76-77)

³²No original: • Is the topic within the students' likely age group experience and level? • Is the topic suitable for every student? The topic should not be biased in any way (in terms of urban/rural, boy/girl, cultural, etc.). • Are there any unsuitable topics such as war, death, politics and religious beliefs that may offend or distress some students? These should be avoided. • Is there too great an assumption of cultural knowledge in the topic? There should not be. • The topic of the text should be unseen but familiar, though not too familiar to the students, i.e., the text should be unseen but students have an existing schema (organized mental framework) for the topic. • Is the style of dialogue/passage familiar to the students? Type of the text should be similar to those in the textbook, e.g., description, narrative, argument. Is the topic content appropriate for the target situation requirements of the students being tested?

3.16 Insumo

Quando falamos de insumo em um teste, referimo-nos às informações que compõem o material, recursos disponíveis nas tarefas de um teste para uso da parte do candidato. A função do insumo é auxiliar a produção de uma resposta apropriada. Um exemplo seria um texto motivador como prévia para performance escrita. A natureza do insumo pode ser verbal, não-verbal ou uma combinação dos seguintes:

- ilustrações de texto ou fotografias (por exemplo, uma imagem de uma sala com pessoas);
- falado (áudio, vídeo, ao vivo);
- texto escrito ou símbolos escritos. (WEIR, 2005, p.77)³³

Se o insumo fornecido reflete a situação-alvo da vida real e garante um equilíbrio entre o conhecimento do conteúdo requerido pela prova e o conhecimento do conteúdo da parte dos candidatos, então, há respaldo teórico para a manutenção de tal insumo em testes de escrita, ou fala.

Para melhor investigar a natureza do insumo de um teste, devemos considerar um conjunto de variáveis linguísticas aplicáveis à tarefa e o quão apropriado esse conjunto pode ser. Para tal, quando falamos de insumo, há de se pensar nos aspectos lexicais, estruturais e funcionais do insumo.

3.16.1 Aspectos Lexicais do Insumo

O insumo lexical refere-se às palavras utilizadas em um enunciado, comando, ou tarefa. Textos com vocábulos mais comuns, ou seja, de alta frequência, tendem a ser mais fáceis do que textos com mais vocabulário de baixa frequência.

Para candidatos cujo nível de inglês é mais baixo (A1/A2 [COUNCIL OF EUROPE,

³³ No original: • text illustrations or photographs (e.g., a picture of a room with people in it); • spoken (audio, video, live); • written text or written symbols.

2001]), precisamos examinar a variedade de formulários de idiomas com os quais tais candidatos seriam capazes de lidar. O texto a servir de material de insumo deve considerar os tipos de conhecimento em relação aos potenciais itens lexicais desconhecidos aos candidatos-alvo. Para alunos de inglês para fins específicos (ESP - *English for Specific Purposes*) de nível superior, precisamos examinar se os itens lexicais são apropriados em termos de vocabulário comum à área, assim como os itens lexicais de cunho técnico e subtécnico (WEIR, 2005).

A pergunta quanto ao léxico de um teste é, em resumo: *os itens lexicais no teste estão no texto de insumo e são exigidos na produção de maneira adequada ao nível dos candidatos?* (WEIR, 2005)

3.16.2 Aspectos Estruturais do Insumo

A familiaridade, ou não-familiaridade, de um candidato com padrões sintáticos, assim como variações de uso em termos de fonética e fonologia, podem influenciar a compreensão da tarefa como um todo. O uso de construções sintáticas longas, complexas, assim como a presença de estilo³⁴ em um enunciado ou comando podem impedir o ouvinte de compreender a linguagem utilizada (ROST, 1990).

Em resumo: *os itens gramaticais no texto estão no texto de insumo e os exigidos como produção (por parte da grade de avaliação) são apropriados para o nível dos candidatos?*

3.16.3 Aspectos Funcionais do Insumo

Função é o termo usado para descrever a característica ilocucionária do que é dito. O termo alude às funções comunicativas, que podem relacionar-se aos atos de persuadir, aconselhar, descrever, dentre outras (SEARLE, 1969; O'SULLIVAN et al; 2002). A validade de um teste é aumentada quanto maior for o número de passagens e itens presentes

³⁴ Alinhamento estilística à situação social que os interactantes estão envolvidos. (CAPPELLA, 1996)

relacionados ao contexto social simulado na tarefa.

Em resumo: *as funções elencadas no teste estão claras no texto de insumo e são exigidas como produção de maneira apropriada para o nível do candidato?*

3.17 Resumo das perguntas advindas do modelo sociocognitivo de validade de contexto

Após a breve apresentação de cada um dos itens presentes em nossa adaptação do modelo de Weir (2005) para uma investigação sociocognitiva da validade de contexto, colocamos cada uma das perguntas que investigaremos e sumarizam a atenção que devemos dar para cada item da prova a seguir.

Quadro 3 - Resumo das perguntas do modelo sociocognitivo de Weir (2005) para investigação de validade de contexto

Assunto	Pergunta
Instruções, enunciados e comandos	<i>[A](s) [instruções, enunciados e] comando(s) do teste são preciso(s) e acessível(is)?</i>
Propósito	<i>O objetivo do teste é inequivocamente claro para o candidato? É um propósito apropriado?</i>
Formato da resposta	<i>Existe alguma evidência de que o formato de resposta do teste pode afetar o desempenho de seu candidato?</i>
Critérios Disponíveis	<i>Os critérios a serem usados para pontuar as tarefas da prova são explícitos para os candidatos e os examinadores?</i>
Peso	<i>De que maneira um item que testa um aspecto X é mais ou menos importante do que um item que testa um aspecto Y? Há alguma ponderação para os diferentes componentes de um teste que tenha sido devidamente justificada?</i>
Ordem	<i>Os itens e as tarefas de um teste estão em uma ordem justificável?</i>
Limitações de tempo	<i>O tempo para cada parte do teste, por exemplo, preparação e conclusão, é apropriado?</i>

Modo de discurso	<i>O modo de discurso é, afinal, apropriado para as habilidades ou estratégias que se deseja avaliar?</i>
Canal de comunicação	<i>O canal é apropriado para os requisitos de situação-alvo dos alunos que estão sendo testados?</i>
Duração	<i>A extensão do texto é apropriada para os requisitos da situação-alvo dos alunos que estão sendo testados?</i>
Natureza da informação	<i>O tipo de informação é apropriado para os requisitos da situação-alvo dos alunos que estão sendo testados?</i>
Conhecimento do conteúdo necessário	<ul style="list-style-type: none"> • <i>O assunto está de acordo com a experiência e o nível provável da faixa etária dos alunos?</i> • <i>O assunto é adequado para todos os alunos? O assunto não deve ser tendencioso de nenhuma forma (em termos de urbano/rural, menino/menina, cultural, etc.).</i> • <i>O assunto do texto deve ser pouco perceptível socialmente, mas familiar; embora não muito familiar para os alunos, ou seja, o texto deve ser pouco perceptível socialmente, mas os alunos têm um esquema existente (estrutura mental organizada) para o assunto.</i> • <i>O estilo de diálogo/passagem é familiar aos alunos? O tipo de texto deve ser semelhante ao do livro didático, por exemplo, descrição, narrativa, argumento.</i> • <i>O conteúdo do assunto é apropriado para os requisitos da situação-alvo dos alunos que estão sendo testados?</i>
Insumo:	<i>Os itens lexicais no teste estão no texto de insumo e são exigidos na produção de maneira adequada ao nível dos candidatos?</i>
Lexical	<i>Os itens gramaticais no texto estão no texto de insumo e os exigidos como produção (por parte da grade de avaliação) são apropriados para o nível dos candidatos?</i>
Estrutural	<i>As funções elencadas no teste estão claras no texto de insumo e são exigidas como produção de maneira apropriada para o nível do candidato?</i>
Funcional	<i>O(s) comando(s) do teste são preciso(s) e acessível(is)?</i>

Fonte: (WEIR, 2005)

Em resumo, cremos que a unificação do modelo sociocognitivo de validação de testes de Weir (2005), somado às perspectivas linguísticas e à teoria sociocultural, elencadas e desenvolvidas no capítulo anterior, nos permitem uma lente de investigação multifocal, que fortalece a análise. Ademais e, sobretudo, as concepções de contexto de Van Dijk (2012),

Matthiessen e Halliday (2009), retomadas por Ramos (2007) e Meurer (2006), fundamentam a discussão sobre esse tema, que é foco da pesquisa. Isso porque as noções de Halliday a respeito de aspectos culturais e situacionais vêm ao encontro do que é entendido no exame como contexto geral e contextos específicos. Tais noções constituem o contexto mental, proposto por Van Dijk (2012), contexto esse, embora não explícito, é depreendido no construto do exame IELTS em relação ao conhecimento que se espera que o candidato demonstre ter. Como os conceitos de contextos de Halliday não são excludentes, mas solidários para constituição do construto mental desenhado por Van Dijk, no nosso entendimento, não há como abordá-los separadamente no âmbito desta pesquisa, que é de fundamento sociocognitivista.. Assim sendo, utilizaremos as concepções de ambos os modelos de contexto, do capítulo referente à base sociocognitiva linguística e da base sociocognitiva avaliativa para investigar os critérios que subjazem a prova oral do IELTS, seus propósitos e o alinhamento de ambos em relação à validade do componente contextual do exame.

No próximo capítulo, apresentamos o desenho do procedimento metodológico que adotamos para nosso trabalho investigativo.

Capítulo 4 - Referencial Metodológico

Este capítulo visa descrever os percursos metodológicos utilizados nesta dissertação a fim de demonstrar os usos que faremos do referencial teórico. Para investigar a validade de contexto das tarefas e da grade de avaliação do exame oral do IELTS em relação aos seus propósitos sob uma perspectiva sociocognitiva, discorreremos, inicialmente, sobre o tipo de pesquisa e suas especificidades, a estrutura da prova e as tarefas selecionadas, as razões subjacentes à escolha do paradigma qualitativo-interpretativista, bem como a técnica e os procedimentos de investigação.

4.1 A Pesquisa Qualitativa

A fim de responder às questões de pesquisa desta dissertação, apoiamo-nos sobre o paradigma qualitativo-interpretativista uma vez que esse nos permite um olhar pessoal, conjectural e científico. Tal paradigma, de maneira geral, viabiliza a utilização de diferentes fontes, métodos, instrumentos e procedimentos para a análise e reflexões quanto às questões de pesquisa, o que nos leva a uma triangulação a ser definida adiante.

Gibbs (2009) assevera que há características comuns à toda pesquisa qualitativa. São elas, cinco: interesse pela experiência em seu contexto original; flexibilidade para reformulação e desenvolvimento de (novas) questões de pesquisa, em oposição às pré-definidas; possibilidade de reformulação da metodologia ante o surgimento de novos paradigmas; valorização da capacidade crítica, da vivência na pesquisa e das contribuições do pesquisador; importância significativa ao contexto em que a pesquisa foi conduzida, associando sempre “quando”, “como”, “onde” e “quem” como pontos relevantes da pesquisa juntamente “daquilo” que se pesquisa. Ao observar o modelo sociocognitivo proposto por Van Dijk (2012) e Weir (2005), é notável que todas as condições elencadas por Gibbs estão presentes, nos alocando, dessa forma, no paradigma qualitativo de investigação. Isso se estende, também, à nossa concepção de língua em nossas bases teóricas. Vê-se, então, que a escolha do paradigma parte tanto da necessidade, quanto da natureza de nossa pesquisa.

Sobre essa última, visamos investigar a validade de construto das tarefas orais e da grade avaliativa do exame IELTS de maneira sociocognitiva, isto é, em relação às maneiras que o instrumento, a sociedade e seus usuários se alinham. A investigação será conduzida, primordialmente, por meio do modelo Sociocognitivo de Weir (2005), apresentado no capítulo anterior. Recorrendo às nossas bases teóricas que advogam em favor do paradigma Sociocognitivo quanto à formação de habilidades linguísticas que, somente por meio de competências interacionais, manifestas no nosso dia-a-dia, que nossas habilidades linguísticas afloram-se e regulam-se. Elas dependem intrinsecamente de nichos sociais, pois cada contexto exige um determinado comportamento verbal, paraverbal e não-verbal que depende de relações sociointeracionais (expressões faciais, corporais, etc.) para que se possa firmar e formar significado da parte do enunciador, seus enunciado, para com seus interactantes.

Assim sendo, não é difícil notar que as perguntas de Gibbs (2009) se relacionam com o parágrafo anterior: o “quando”, “como”, “onde”, “por quê”, “por quem” são voltadas à avaliação e à validade do exame oral do IELTS. Há de se considerar também que, uma vez que a TSC não provê uma corrente linguística, mas sim uma teoria psicológica sobre a formação do pensamento e da sociedade humana, o apelo à Sociocognição e suas vertentes linguísticas se faz necessário e oportuno, uma vez que essa, segundo Tomasello (2003), adequa-se e alinha-se à TSC, complementando-a quanto à compreensão do mundo. Assim, podemos partir para observações focais, voltadas para a alteridade, para o “olhar do outro”. A perspectiva sociocognitiva, então, amplia o escopo de análise de nossa pesquisa para além do falar, para além da linguística, e nos leva ao nível social e sociointeracional da produção de(a) linguagem.

4.2 Modalidades do exame

Apresentaremos um pouco da história desse exame para que possamos melhor discutí-lo. Diferentemente do CPE, citado na introdução de nosso trabalho, que demonstrava raízes arraigadas no Método Tradicional (*Grammar Translation Method*) de ensino de línguas, as versões anteriores do IELTS já contavam com diferenças significantes em relação ao CPE. Inicialmente, o *English Proficiency Test Battery* (EPTB), teste que surgiu em 1963,

examinava somente as habilidades de leitura e compreensão oral de seus candidatos. Em 1980, junto das teorias de Hymes (1972) e o surgimento de cursos de inglês para fins específicos, surge o *English Language Testing Service* (ELTS), que substituiu o EPTB, a fim de tornar os testes mais práticos e relevantes³⁵. O ELTS possuía seis módulos, cinco deles eram específicos (ciências biológicas, sociais, naturais, tecnológicas e médicas) sendo o sexto um módulo acadêmico geral (MANHATTAN, 2021).

O formato do IELTS, quase como conhecemos hoje, surge somente em 1989, com quatro módulos individuais (fala, leitura, audição e escrita). A junção das organizações British Council (BC), *International Development Program: IELTS Australia* (IDP) e a *Cambridge Assessment English* (CAE), galgaram, por influência do IDP, um alcance internacional anteriormente inatingível. Os módulos de audição e fala eram gerais, tanto para a modalidade acadêmica, quanto para a modalidade de treinamento geral. Já as provas de leitura e escrita da modalidade acadêmica procuravam assemelhar-se com o curso pretendido do candidato, sendo elas divididas em três seções: tecnologia, ciências biológicas e ciências sociais.

Atualmente, o IELTS conta com os mesmos quatro módulos, porém, a parte acadêmica de leitura não se volta para módulos específicos curriculares, alteração que também foi aplicada à prova de escrita. O teste não deixa de apresentar enfoque acadêmico em sua respectiva modalidade, uma vez que há a presença, em um nível geral, de textos tipicamente mais longos, assim como a exigência de que os alunos saibam interpretar gráficos. Já na modalidade de treinamento geral, outros gêneros discursivos são exigidos dos alunos, gêneros estes voltados para funções sociais, tal como a leitura de placas, de anúncios, dentre outros; enquanto na prova escrita, prioriza-se, por exemplo, uma carta (in)formal, de agradecimento ou reclamação, em vez da interpretação de um gráfico ou mapa (IELTS, 2021).

O exame IELTS, como um todo, é regido por seus próprios documentos e grades de avaliação. Porém, pode ser utilizado junto ao QEQR (COUNCIL OF EUROPE, 2001) para uma melhor compreensão dos níveis de proficiência que a prova se propõe a avaliar. Tais níveis variam de 1 a 9, certificando somente os candidatos que atingem de 4 a 9. No entanto, todos os níveis do IELTS podem ser organizados de acordo com os do QEQR, que variam de A1 ao C2. cremos que isso se dá em decorrência dos esforços históricos, desde a década de

³⁵ Apesar de compor o corpo do texto, a palavra “relevante” não reflete nossa opinião sobre quaisquer aspectos da realidade. Relevância, tal como Bachman (1990) aponta, é um construto extremamente volátil; assim sendo, cremos que a relevância referida pela nossa fonte original referia-se ao fim da prova, neste caso, universidades.

1970, para que exames e padrões fossem estabelecidos na União Européia (PICCARDO; NORTH, 2019).

A menção ao QEER (COUNCIL OF EUROPE, 2001) visa esclarecer que é possível traçar relação entre a própria grade de avaliação do IELTS e a de outros sistemas, mas devemos ter em mente que as relações entre uma grade de avaliação e outra não podem ser interpretadas como biunívocas, uma vez que a operacionalização dos construtos em uma prova de proficiência tem em vista propósitos específicos, diferentemente do que é descrito num documento como o QEER (COUNCIL OF EUROPE, 2001), que é de cunho psicopedagógico. Nossa pesquisa tem como paradigma a perspectiva sociocognitiva de língua, que não sabemos, em princípio, o quão alinhada está às concepções de língua tomadas por essa prova.

Apontamos que, apesar de termos acesso às grades de avaliação do IELTS, não temos permissão para utilizá-las. Os documentos oficiais que regem a prova são confidenciais e, portanto, nos ateremos às versões disponíveis nas publicações presentes no próprio *website* do exame IELTS e os demais materiais livremente divulgados.

O IELTS contempla duas modalidades: a acadêmica e a de treinamento geral. Para explicar a estrutura geral da prova e as diferenças entre ambas as modalidades, trouxemos informações sobre sua estrutura, intenção das questões e descritores. Começamos pela divisão da prova, que é feita em quatro etapas: o teste de fala (*speaking*); audição (*listening*); leitura (*reading*) e escrita (*writing*); a organização dos testes é feita da seguinte maneira:

Quadro 4 - Panorama do exame IELTS 1

Habilidades Produtivas	Fala	Escrita
Tempo de prova	11 - 14 minutos	60 minutos - 20 para a primeira tarefa; 40 para a segunda
Questões	Variável	2 questões
Partes Acadêmico	3 partes 1ª - Perguntas sobre situações familiares 2ª - Perguntas sobre um tópico qualquer escolhido pelo	Tarefa 1 - análise de gráfico, mapa ou tabela Tarefa 2 - texto discursivo-argumentativo

	examinador 3ª - Perguntas que desenvolvam o tópico da parte 2 de maneira mais geral e abstrata	
Partes Treinamento geral	Sem alterações	Tarefa 1 - carta formal/semi formal/informal Tarefa 2 - texto discursivo-argumentativo

Fonte: IELTS (2021)

Quadro 5 - Panorama do exame IELTS 2

Habilidades Receptivas	Audição	Leitura
Tempo de prova	40 minutos totais - 30 para o exercício e 10 para o preenchimento	60 minutos - 50 minutos para o teste e 10 para o preenchimento
Questões	40 questões	40 questões
	4 partes 1ª - diálogo cotidiano entre duas pessoas 2ª - monólogo de cunho variado - como sobre as instalações locais - sobre aspectos sociais 3ª - diálogo universitário entre duas pessoas 4ª - monólogo sobre um assunto acadêmico	3 textos de diferentes gêneros relacionados à leitura acadêmica.
Partes Treinamento geral	Sem alterações	3 partes - uma social, uma contextual e uma geral.

Fonte: IELTS (2021)

Uma vez que nosso objetivo é investigar a prova oral do teste IELTS a fim de observar o que ela revela sobre a natureza sociointeracional do exame, detalhamos um pouco mais sobre sua composição. A seção oral do exame é organizada da seguinte forma (SEEDHOUSE et al, 2004).

Parte 1 - o examinador faz perguntas sobre o candidato e diversos tópicos familiares, tais como seu lar, família, trabalho, estudos e interesses. Esta parte dura entre quatro e cinco minutos. (IELTS, 2021)

Parte 2 - o candidato recebe um cartão contendo um tópico, sobre o qual deverá discorrer sobre. Os examinandos têm até 1 minuto para se preparar. Ao fim dos dois minutos de preparação, o examinador pedirá ao candidato que comece a discorrer sobre o assunto por até dois minutos. Uma vez que o tempo acaba, ao menos uma, ou até duas perguntas, podem ser feitas pelo examinador. Esta parte varia entre três e cinco minutos. (IELTS, 2021)

Parte 3 - o examinando recebe mais perguntas relacionadas aos temas da parte 2. Estas perguntas extras aparentam ter como objetivo discutir os mesmos tópicos sob uma ótica mais generalizada. A parte 3 dura entre três e cinco minutos. (IELTS, 2021)

As perguntas da prova, retiradas do artigo de Seedhouse e Eggbert (2004), foram transcritas a seguir e serão de onde partiremos para iniciar nossa investigação das tarefas do exame oral do IELTS. Essas tarefas são advindas de dois diferentes exames e, logo, de dois candidatos. Explicamos que optamos por duas transcrições em vez de uma a fim de angariar mais amostras do exame.

Creemos que, a partir de uma amostra do exame oral com mais tarefas, seremos capazes de conduzir uma investigação de cunho sociocognitivo voltada aos diversos contextos relativos ao construto de oralidade e também os diversos tipos de validade a serem investigados. Partimos do pressuposto de Luoma (2004) que as questões da amostra podem ter funções generalizadas em outros temas para viabilizar mais análises de nossos excertos.

Quadro 6 - Transcrição e tradução das amostras do exame

Original	Tradução
Part 1	Parte 1
Candidate 1	Candidato 1

<p>Introduction</p> <p>1 - Can you tell me where you are from?</p> <p>2 - Can we talk about where you live? Could you describe where you live?</p> <p>3 - What do you like about living there?</p> <p>4 - Are there things you don't like about it?</p> <p>5 - How could you improve the city?</p> <p>Let's move on to the topic of food and restaurants</p> <p>1 - What kind of food do you like to eat?</p> <p>2 - Is there any food you don't like?</p> <p>3 - What are some of the advantages and disadvantages of eating in a restaurant?</p> <p>4 - What's the good thing about eating in a restaurant?</p> <p>Let's talk about films?</p> <p>1 - Do you enjoy watching films?</p> <p>2 - How often do you watch films?</p>	<p>Introdução</p> <p>1 - Você pode me dizer de onde você é?</p> <p>2 - Podemos conversar sobre onde você mora? Você poderia descrever onde você mora?</p> <p>3 - Do que você gosta de morar lá?</p> <p>4 - Há coisas que você não goste?</p> <p>5 - Como você poderia melhorar a cidade?</p> <p>Vamos passar para o tópico de alimentação e restaurantes</p> <p>1 - Que tipo de comida você gosta de comer?</p> <p>2 - Existe alguma comida que você não goste?</p> <p>3 - Quais são algumas das vantagens e desvantagens de comer em um restaurante?</p> <p>4 - Qual é a vantagem de comer em um restaurante?</p> <p>Vamos falar sobre filmes?</p> <p>1 - Você gosta de assistir filmes?</p> <p>2 - Com que frequência você assiste filmes?</p>
<p>Candidate 2</p>	<p>Candidato 2</p>
<p>Introduction</p> <p>1 - Tell me about your studies.</p> <p>2 - What do you like most about your studies?</p> <p>3 - Are there any things you don't like about your studies?</p> <p>4 - What qualifications or certificates do you hope to get?</p> <p>Let's talk about activities you enjoy in your free time.</p> <p>1 - When do you have free time?</p> <p>2 - What free time activities are most popular where you live?</p> <p>3 - How important is free time in people's lives?</p> <p>Can we talk about your childhood?</p> <p>1 - Where did you grow up?</p> <p>2 - Was it a good place for children? Why?</p> <p>3 - Where did you usually play?</p>	<p>Introdução</p> <p>1 - Conte-me sobre seus estudos.</p> <p>2 - O que você mais gosta em seus estudos?</p> <p>3 - Há alguma coisa que você não goste em seus estudos?</p> <p>4 - Quais qualificações ou certificados você espera obter?</p> <p>Vamos falar sobre as atividades que você gosta no seu tempo livre.</p> <p>1 - Quando você tem tempo livre?</p> <p>2 - Quais são as atividades de lazer mais populares onde você mora?</p> <p>3 - Qual a importância do tempo livre na vida das pessoas?</p> <p>Podemos falar sobre sua infância?</p> <p>1 - Onde você cresceu?</p> <p>2 - Era um bom lugar para crianças? Por quê?</p>

<p>4 - Do you think childhood is different today than when you were a child?</p>	<p>3 - Onde você costumava jogar? 4 - Você acha que a infância hoje é diferente de quando você era criança?</p>
<p>Part 2</p>	<p>Parte 2</p>
<p>Now let's move on to the next part and I am going to give you a topic, and I'd like you to talk about it for one or two minutes. Before you talk you will have one minute to think about what you are going to say. And you can make some notes if you wish. All right? Do you understand?</p> <p>Here's some paper for making notes and here's your topic. I'd like you to describe a trip that you once went on.³⁶</p> <p>(written on task card)</p> <p>1 - Where you went 2 - Why you went there 3 - How you traveled 4 - What you enjoyed most 5 - If you would repeat this trip and why Do you generally enjoy traveling?</p>	<p>Agora, vamos passar para a próxima parte e vou lhe dar um tópico, e gostaria que você falasse sobre isso por um ou dois minutos. Antes de falar, você terá um minuto para pensar no que vai dizer. E você pode fazer algumas anotações, se desejar. Tudo bem? Você entende?</p> <p>Aqui está um papel para fazer anotações e este é o seu tópico. Eu gostaria que você descrevesse uma viagem que você fez uma vez.</p> <p>(escrito no cartão de tarefas)</p> <p>1 - Para onde você foi 2 - Por que você foi lá 3 - Como você viajou 4 - O que você mais gostou 5 - Se você repetiria esta viagem e por quê Você geralmente gosta de viajar?</p>
<p>Part 3</p>	<p>Parte 3</p>
<p>We've been talking about a trip that you went on and I'd like to discuss with you one or two more general questions related to this..</p> <p>Tópico 1 - Let's think about travel and transport.</p> <p>What's the most popular way to travel long distance in your country?</p> <p>How do you think we're going to travel in the future?</p> <p>Tópico 2 - Let's talk about public and private transportation.</p> <p>Can you describe the public transport system in your country?</p>	<p>Estamos falando sobre uma viagem que você fez e gostaria de discutir com você uma ou duas questões gerais relacionadas a isso.</p> <p>Tópico 1 - Vamos pensar sobre viagens e transporte.</p> <p>Qual é a forma mais popular de viajar de longa distância em seu país?</p> <p>Como você acha que vamos viajar no futuro?</p> <p>Tópico 2 - Vamos falar sobre transporte público e privado.</p>

³⁶ Inferimos as perguntas a partir das respostas dos candidatos examinados no artigo base. Ambos receberam a mesma parte 2 e 3, apresentando variações somente na parte 1.

Can you speculate on any measures that will be taken to reduce pollution in the future?	<p>Você pode descrever o sistema de transporte público em seu país?</p> <p>Você pode especular sobre as medidas que serão tomadas para reduzir a poluição no futuro?</p>
---	--

Fonte: (SEEDHOUSE; EGGBERT, 2004)

Paralelamente às tarefas, temos a grade de avaliação. A grade de avaliação que transpomos é advinda do artigo de Seedhouse, Harris, Naeb e Unstunel (2014). A grade de avaliação é o documento utilizado como referência para aferir o desempenho dos candidatos no exame oral do IELTS. Divida em quatro aspectos, alcance gramatical, fluência e coerência, pronúncia e recurso lexical, a referida grade de avaliação possui subdivisões, como já constatado, que variam de 1 a 9 para cada um dos referidos critérios, sendo cada critério dividido de maneira diferente. Adentraremos a grade de avaliação em mais detalhes no capítulo de análise. Nos parece que as três partes das entrevistas são analisadas como um *continuum* quanto ao desempenho do candidato, comparado-o e contrastando-o com os critérios elencados na grade avaliativa. A relação entre as tarefas e os critérios é parte de nossa análise, bem como os fins profissionais e acadêmicos envolvidos nas diferentes modalidades do exame oral do IELTS.

Quadro 7 - Grade de Avaliação do IELTS completa (com tradução)

Fluency and coherence	Fluência e coerência
<p style="text-align: center;">9</p> <ul style="list-style-type: none"> •speaks fluently with only rare repetition or self-correction; •any hesitation is content-related rather than to find words or grammar •speaks coherently with fully appropriate cohesive features •develops topics fully and appropriately 	<p style="text-align: center;">9</p> <ul style="list-style-type: none"> • fala fluentemente apenas com raras repetições ou autocorreção; • qualquer hesitação está relacionada ao conteúdo, e não para encontrar palavras ou gramática • fala coerentemente com recursos coesivos totalmente apropriados • desenvolve tópicos de forma completa e adequada
<p style="text-align: center;">8</p> <ul style="list-style-type: none"> •speaks fluently with only occasional repetition or self-correction; hesitation is usually content-related and only rarely to search for language •develops topics coherently and appropriately 	<p style="text-align: center;">8</p> <ul style="list-style-type: none"> • fala fluentemente apenas com repetição ou autocorreção ocasional; a hesitação geralmente está relacionada ao conteúdo e apenas raramente à pesquisa de idioma • desenvolve tópicos de forma coerente e adequada
<p style="text-align: center;">7</p> <ul style="list-style-type: none"> •speaks at length without noticeable effort or loss of coherence 	<p style="text-align: center;">7</p> <ul style="list-style-type: none"> • fala longamente sem esforço perceptível ou perda de coerência

<ul style="list-style-type: none"> •may demonstrate language-related hesitation at times, or some repetition and/or self-correction •uses a range of connectives and discourse markers with some flexibility 	<ul style="list-style-type: none"> • pode demonstrar hesitação relacionada à linguagem às vezes, ou alguma repetição e / ou autocorreção • usa uma variedade de conectivos e marcadores de discurso com alguma flexibilidade
<p style="text-align: center;">6</p> <ul style="list-style-type: none"> •is willing to speak at length, though may lose coherence at times due to occasional repetition, self-correction or hesitation •uses a range of connectives and discourse markers but not always appropriately 	<p style="text-align: center;">6</p> <ul style="list-style-type: none"> • está disposto a falar longamente, embora possa perder coerência às vezes devido à repetição ocasional, autocorreção ou hesitação • usa uma variedade de conectivos e marcadores de discurso, mas nem sempre de forma adequada
<p style="text-align: center;">5</p> <ul style="list-style-type: none"> •usually maintains flow of speech but uses repetition, self correction and/or slow speech to keep going •may over-use certain connectives and discourse markers •produces simple speech fluently, but more complex communication causes fluency problems 	<p style="text-align: center;">5</p> <ul style="list-style-type: none"> • geralmente mantém o fluxo da fala, mas usa repetição, autocorreção e/ou fala lenta para continuar • pode usar excessivamente certos conectivos e marcadores de discurso • produz um discurso simples com fluência, mas uma comunicação mais complexa causa problemas de fluência
<p style="text-align: center;">4</p> <ul style="list-style-type: none"> •cannot respond without noticeable pauses and may speak slowly, with frequent repetition and self-correction •links basic sentences but with repetitious use of simple connectives and some breakdowns in coherence 	<p style="text-align: center;">4</p> <ul style="list-style-type: none"> • não é capaz de responder sem pausas perceptíveis e por vezes fala devagar, com repetição frequente e autocorreção • vincula frases básicas, mas com o uso repetitivo de conectivos simples e algumas falhas na coerência
<p style="text-align: center;">3</p> <ul style="list-style-type: none"> •speaks with long pauses •has limited ability to link simple sentences •gives only simple responses and is frequently unable to convey basic message 	<p style="text-align: center;">3</p> <ul style="list-style-type: none"> • fala com longas pausas • tem capacidade limitada de vincular frases simples • dá apenas respostas simples e frequentemente é incapaz de transmitir a mensagem básica
<p style="text-align: center;">2</p> <ul style="list-style-type: none"> •pauses lengthily before most words •little communication possible 	<p style="text-align: center;">2</p> <ul style="list-style-type: none"> • faz uma longa pausa antes da maioria das palavras • pouca comunicação possível
<p style="text-align: center;">1</p> <ul style="list-style-type: none"> •no communication possible •no rateable language 	<p style="text-align: center;">1</p> <ul style="list-style-type: none"> • nenhuma comunicação possível • sem linguagem classificável
<p style="text-align: center;">0</p> <ul style="list-style-type: none"> •does not attend 	<p style="text-align: center;">0</p> <ul style="list-style-type: none"> • não comparece
Pronunciation	Pronúncia
<p style="text-align: center;">9</p> <ul style="list-style-type: none"> •uses a full range of pronunciation features with precision and subtlety •sustains flexible use of features throughout •is effortless to understand 	<p style="text-align: center;">9</p> <ul style="list-style-type: none"> • usa uma gama completa de recursos de pronúncia com precisão e sutileza • mantém o uso flexível de recursos • é fácil de entender

<p style="text-align: center;">8</p> <ul style="list-style-type: none"> •uses a wide range of pronunciation features •sustains flexible use of features, with only occasional lapses •is easy to understand throughout; L1 accent has minimal effect on intelligibility 	<p style="text-align: center;">8</p> <ul style="list-style-type: none"> • usa uma ampla gama de recursos de pronúncia • mantém o uso flexível de recursos da fala conectada, com apenas lapsos ocasionais • é fácil de entender durante a fala; o sotaque da L1 tem efeito mínimo na inteligibilidade
<p style="text-align: center;">7</p> <ul style="list-style-type: none"> •shows all the positive features of Band 6 and some, but not all, of the positive features of Band 8 	<p style="text-align: center;">7</p> <ul style="list-style-type: none"> • mostra todas as características positivas da Banda 6 e algumas, mas não todas, das características positivas da Banda 8
<p style="text-align: center;">6</p> <ul style="list-style-type: none"> •uses a range of pronunciation features with mixed control •shows some effective use of features but this is not sustained •can generally be understood throughout, though mispronunciation of individual words or sounds reduces clarity at times 	<p style="text-align: center;">6</p> <ul style="list-style-type: none"> • usa uma variedade de recursos de pronúncia com controle misto • mostra algum uso eficaz de recursos, mas isso não é sustentado • geralmente pode ser totalmente compreendido, embora a pronúncia incorreta de palavras ou sons individuais reduza a clareza às vezes
<p style="text-align: center;">5</p> <ul style="list-style-type: none"> •shows all the positive features of Band 4 and some, but not all, of the positive features of Band 6 	<p style="text-align: center;">5</p> <ul style="list-style-type: none"> • mostra todas as características positivas da Banda 4 e algumas, mas não todas, das características positivas da Banda 6
<p style="text-align: center;">4</p> <ul style="list-style-type: none"> •uses a limited range of pronunciation features •attempts to control features but lapses are frequent •mispronunciations are frequent and cause some difficulty for the listener 	<p style="text-align: center;">4</p> <ul style="list-style-type: none"> • usa uma gama limitada de recursos de pronúncia • tentativas de controlar recursos, mas lapsos são frequentes • erros de pronúncia são frequentes e causam alguma dificuldade para o ouvinte
<p style="text-align: center;">3</p> <ul style="list-style-type: none"> •shows some of the features of Band 2 and some, but not all, of the positive features of Band 4 	<p style="text-align: center;">3</p> <ul style="list-style-type: none"> • mostra algumas das características da Banda 2 e algumas, mas não todas, as características positivas da Banda 4
<p style="text-align: center;">2</p> <ul style="list-style-type: none"> •Speech is often unintelligible 	<p style="text-align: center;">2³⁷</p> <ul style="list-style-type: none"> • A fala muitas vezes é ininteligível
Lexical Resource	Recurso Lexical
<p style="text-align: center;">9</p> <ul style="list-style-type: none"> •uses vocabulary with full flexibility and precision in all topics •uses idiomatic language naturally and accurately 	<p style="text-align: center;">9</p> <ul style="list-style-type: none"> • usa vocabulário com total flexibilidade e precisão em todos os tópicos • usa linguagem idiomática com naturalidade e precisão
<p style="text-align: center;">8</p> <ul style="list-style-type: none"> •uses a wide vocabulary resource readily and flexibly to convey precise meaning •uses less common and idiomatic vocabulary skilfully, with occasional inaccuracies •uses paraphrase effectively as required 	<p style="text-align: center;">8</p> <ul style="list-style-type: none"> • usa um amplo recurso de vocabulário de forma rápida e flexível para transmitir um significado preciso • usa vocabulário menos comum e idiomático

³⁷ Não há descritores presentes para os níveis 0 e 1 nos quesitos “recurso lexical”, “alcance gramatical” e “pronúncia”.

	<p>habilmente, com imprecisões ocasionais</p> <ul style="list-style-type: none"> • usa a paráfrase de forma eficaz, conforme necessário
<p>7</p> <ul style="list-style-type: none"> •uses vocabulary resource flexibly to discuss a variety of topics •uses some less common and idiomatic vocabulary and shows some awareness of style and collocation, with some inappropriate choices •uses paraphrase effectively 	<p>7</p> <ul style="list-style-type: none"> • usa recursos de vocabulário de forma flexível para discutir uma variedade de tópicos • usa algum vocabulário menos comum e idiomático e mostra alguma consciência de estilo e colocação, com algumas escolhas inadequadas • usa a paráfrase de forma eficaz
<p>6</p> <ul style="list-style-type: none"> •has a wide enough vocabulary to discuss topics at length and make meaning clear in spite of inappropriacies •generally paraphrases successfully 	<p>6</p> <ul style="list-style-type: none"> • tem um vocabulário amplo o suficiente para discutir tópicos extensamente e tornar o significado claro, mesmo que com inadequações • geralmente parafraseia com sucesso
<p>5</p> <ul style="list-style-type: none"> •manages to talk about familiar and unfamiliar topics but uses vocabulary with limited flexibility •attempts to use paraphrase but with mixed success 	<p>5</p> <ul style="list-style-type: none"> • consegue falar sobre tópicos familiares e desconhecidos, mas usa vocabulário com flexibilidade limitada • tentativas de usar a paráfrase, mas com sucesso misto
<p>4</p> <ul style="list-style-type: none"> •is able to talk about familiar topics but can only convey basic meaning on unfamiliar topics and makes frequent errors in word choice •rarely attempts paraphrase 	<p>4</p> <ul style="list-style-type: none"> • é capaz de falar sobre tópicos familiares, mas só consegue transmitir o significado básico de tópicos desconhecidos e comete erros frequentes na escolha das palavras • raramente tenta parafrasear
<p>3</p> <ul style="list-style-type: none"> •uses simple vocabulary to convey personal information •has insufficient vocabulary for less familiar topics 	<p>3</p> <ul style="list-style-type: none"> • usa vocabulário simples para transmitir informações pessoais • tem vocabulário insuficiente para tópicos menos familiares
<p>2</p> <ul style="list-style-type: none"> •only produces isolated words or memorised utterances 	<p>2</p> <ul style="list-style-type: none"> • produz apenas palavras isoladas ou enunciados memorizados
Grammatical range and accuracy	Alcance gramatical e precisão
<p>9</p> <ul style="list-style-type: none"> •uses a full range of structures naturally and appropriately •produces consistently accurate structures apart from 'slips' characteristic of native speaker speech 	<p>9</p> <ul style="list-style-type: none"> • usa uma gama completa de estruturas de forma natural e adequada • produz estruturas consistentemente precisas além das características de "deslizes" da fala do falante nativo
<p>8</p> <ul style="list-style-type: none"> •uses a wide range of structures flexibly •produces a majority of error-free sentences with only very occasional inappropriacies or basic/non-systematic errors 	<p>8</p> <ul style="list-style-type: none"> • usa uma ampla gama de estruturas de forma flexível • produz a maioria das sentenças livres de erros, apenas com inadequações muito ocasionais ou erros básicos/não sistemáticos

7	7
<ul style="list-style-type: none"> •uses a range of complex structures with some flexibility •frequently produces error-free sentences, though some grammatical mistakes persist 	<ul style="list-style-type: none"> • usa uma variedade de estruturas complexas com alguma flexibilidade • frequentemente produz frases sem erros, embora alguns erros gramaticais persistam
6	6
<ul style="list-style-type: none"> •uses a mix of simple and complex structures, but with limited flexibility •may make frequent mistakes with complex structures though these rarely cause comprehension problems 	<ul style="list-style-type: none"> • usa uma mistura de estruturas simples e complexas, mas com flexibilidade limitada • pode cometer erros frequentes com estruturas complexas, embora raramente causem problemas de compreensão
5	5
<ul style="list-style-type: none"> •produces basic sentence forms with reasonable accuracy •uses a limited range of more complex structures, but these usually contain errors and may cause some comprehension problems 	<ul style="list-style-type: none"> • produz frases básicas com precisão razoável • usa uma gama limitada de estruturas mais complexas, mas geralmente contêm erros e podem causar alguns problemas de compreensão
4	4
<ul style="list-style-type: none"> •produces basic sentence forms and some correct simple sentences but subordinate structures are rare •errors are frequent and may lead to misunderstanding 	<ul style="list-style-type: none"> • produz formas de frases básicas e algumas frases simples corretas, mas estruturas subordinadas são raras • erros são frequentes e podem levar a mal-entendidos
3	3
<ul style="list-style-type: none"> •attempts basic sentence forms but with limited success, or relies on apparently memorised utterances •makes numerous errors except in memorised expressions 	<ul style="list-style-type: none"> • tenta formas de frases básicas, mas com sucesso limitado, ou depende de declarações aparentemente memorizadas • comete vários erros, exceto em expressões memorizadas
2	2
<ul style="list-style-type: none"> •cannot produce basic sentence forms 	<ul style="list-style-type: none"> • não consegue produzir formas de frases básicas

Fonte: (SEEDHOUSE et al 2014)

Assim, o escore final do IELTS é calculado com base na soma, seguido da média ponderada, excluindo resultados .25 e .75 (com base na distribuição de escores presente no *website* do IELTS [2021]) do desempenho do candidato. A competência é indicada de acordo com a nota obtida. Para melhor ilustrar as relações entre notas e desempenho/proficiência, apresentamos um quadro 8, comparativo dos descritores do QECR (COUNCIL OF EUROPE, 2001) e do IELTS (2021).

Quadro 8 - Relações entre os níveis do CEFR e a escala gradativa do IELTS

Divisões de níveis de proficiência do QECR	IELTS' level descriptors	Tradução
C2	Grade: 9 Expert user	Nota: 9 usuário especialista
É capaz de compreender, sem esforço, praticamente tudo o que ouve ou lê. É capaz de resumir as informações recolhidas em diversas fontes orais e escritas, reconstruindo argumentos e fatos de um modo coerente. É capaz de se exprimir espontaneamente, de modo fluente e com exatidão, sendo capaz de distinguir finas variações de significado em situações complexas.	The test taker has fully operational command of the language. Their use of English is appropriate, accurate and fluent, and shows complete understanding.	O candidato possui total domínio operacional do idioma. O uso do inglês é apropriado, preciso e fluente, demonstrando entendimento completo.
C1/C2	Grade: 8 very good user	Nota: 8 usuário muito bom
Apresenta características do nível C1 e algumas, mas não todas, do C2.	The test taker has fully operational command of the language with only occasional unsystematic inaccuracies and inappropriate usage. They may misunderstand some things in unfamiliar situations. They handle complex and detailed argumentation well.	O candidato tem o comando totalmente operacional do idioma, apenas com imprecisões assistemáticas ocasionais e uso impróprio. Eles podem interpretar mal algumas coisas em situações desconhecidas. Eles lidam bem com argumentação complexa e detalhada.
C1	Grade: 7 good user	Nota: 7 bom usuário
É capaz de compreender um vasto número de textos longos e exigentes, reconhecendo os seus significados implícitos. É capaz de se exprimir de forma fluente e espontânea sem precisar procurar muito as palavras. É capaz de usar a língua de modo flexível e eficaz para fins sociais, acadêmicos e profissionais. Pode exprimir-se sobre temas complexos, de forma clara e bem estruturada, manifestando o domínio de mecanismos de organização, de articulação e de coesão do discurso.	The test taker has operational command of the language, though with occasional inaccuracies, inappropriate usage and misunderstandings in some situations. They generally handle complex language well and understand detailed reasoning.	O candidato tem o domínio operacional do idioma, embora com imprecisões ocasionais, uso impróprio e mal-entendidos em algumas situações. Eles geralmente lidam bem com linguagens complexas e entendem o raciocínio detalhado.
B2	Grade: 6 competent user	Nota: 6 usuário competente
É capaz de compreender as ideias principais em textos complexos sobre assuntos concretos e abstratos, incluindo discussões técnicas na sua área de especialidade. É capaz de comunicar com certo grau de	The test taker has an effective command of the language despite some inaccuracies, inappropriate usage and misunderstandings. They can use and understand fairly complex language,	O candidato tem um domínio eficaz do idioma, apesar de algumas imprecisões, uso impróprio e mal-entendidos. Eles podem usar e compreender uma linguagem bastante

espontaneidade com falantes nativos, sem que haja tensão de parte a parte. É capaz de exprimir-se de modo claro e pormenorizado sobre uma grande variedade de temas e explicar um ponto de vista sobre um tema da atualidade, expondo as vantagens e os inconvenientes de várias possibilidades.	particularly in familiar situations.	complexa, especialmente em situações familiares.
B1	Grade: 5 modest user	Nota: 5 usuário aceitável
É capaz de compreender as questões principais, quando é usada uma linguagem clara e estandardizada e os assuntos lhe são familiares (temas abordados no trabalho, na escola e nos momentos de lazer, etc.). É capaz de lidar com a maioria das situações encontradas na região onde se fala a língua-alvo. É capaz de produzir um discurso simples e coerente sobre assuntos que lhe são familiares ou de interesse pessoal. Pode descrever experiências e eventos, sonhos, esperanças e ambições, bem como expor brevemente razões e justificações para uma opinião ou um projeto.	The test taker has a partial command of the language and copes with overall meaning in most situations, although they are likely to make many mistakes. They should be able to handle basic communication in their own field.	O candidato tem um domínio parcial do idioma e lida com o significado geral do contexto na maioria das situações, embora seja provável que vá cometer muitos erros. Eles provavelmente são capazes de lidar com a comunicação básica em seu próprio campo profissional ou acadêmico.
A2	Grade: 4 Modest User	Nota: 4 usuário modesto
É capaz de compreender frases isoladas e expressões frequentes relacionadas com áreas de prioridade imediata (ex.: informações pessoais e familiares simples, compras, meio circundante). É capaz de comunicar em tarefas simples e em rotinas que exigem apenas uma troca de informação simples e direta sobre assuntos que lhe são familiares e habituais. Pode descrever de modo simples a sua formação, o meio circundante e, ainda, referir assuntos relacionados com necessidades imediatas.	The test taker's basic competence is limited to familiar situations. They frequently show problems in understanding and expression. They are not able to use complex language	A competência básica do candidato é limitada a situações familiares. Eles frequentemente mostram problemas de compreensão e expressão. Eles não são capazes de usar uma linguagem complexa
A1/A2	Grade: 3 Extremely Limited User	Nota: 3 Usuário extremamente limitado
Apresenta características do nível A1 e algumas, mas não todas, do A2.	The test taker conveys and understands only general meaning in very familiar situations. There are frequent breakdowns in communication.	O candidato transmite e compreende apenas o significado geral em situações muito familiares. Existem falhas frequentes na comunicação

A1	Grade: 2 Intermittent user	Nota: 2 usuário intermitente
É capaz de compreender e usar expressões familiares e cotidianas, assim como enunciados muito simples, que visam satisfazer necessidades concretas. Pode apresentar-se e apresentar outros e é capaz de fazer perguntas e dar respostas sobre aspectos pessoais como, por exemplo, o local onde vive, as pessoas que conhece e as coisas que tem. Pode comunicar de modo simples, se o interlocutor falar lenta e distintamente e se mostrar cooperante.	The test taker has great difficulty understanding spoken and written English.	O candidato tem grande dificuldade em entender o inglês falado e escrito.
Pré-A1	Grade: 1 Non-user	Nota: 1 não usuário
	The test taker has no ability to use the language except a few isolated words.	O candidato não consegue usar o idioma, exceto algumas palavras isoladas.
-	Grade: 0 Did not attempt the test	Nota: 0 Não compareceu/fez o teste
-	The test taker did not answer the questions	O participante do teste não respondeu às perguntas

Fonte: IELTS (2021); Council of Europe (2021); Conselho Britânico (2021)

4.3 A pesquisa documental e o interpretativismo

O tipo de pesquisa que fazemos insere-se no escopo, também, das pesquisas documentais. Eis, a frente, definições de pesquisas documentais: “[são] restrita[s] a documentos, escritos ou não” (MARCONI.; LAKATOS, 2003, p. 174) e são “dispersas, sem tratamento analítico, tais como: tabelas estatísticas, jornais, revistas, relatórios, documentos oficiais, cartas, filmes, fotografias, pinturas, tapeçarias, relatórios de empresas, vídeos de programas de televisão, etc.” (FONSECA, 2002, p. 32)”. Isso significa que, com base em nosso referencial, faremos uma coleta de dados a fim de tecer contribuições para nossa investigação. Para Cresswell, a interpretação é “o [momento em que o] pesquisador recua e dá algum significado maior sobre o fenômeno baseado em pontos de vista pessoais, comparações com estudos anteriores ou ambos” (2002, p. 257).

No entanto, Stake (2011) considera que, ao passo que podemos utilizar a liberdade

interpretativa como trunfo investigativo, devemos estar cientes de que quanto maior o grau de subjetividade de uma pesquisa, maior a possibilidade de seu descrédito. O próprio autor (STAKE, 2011) aponta, todavia, que a visão de mundo que concebemos é invariavelmente o resultado do confronto da base de nossos conceitos, isto é, nosso referencial para interpretar o mundo, em relação às nossas próprias experiências, ou seja, o que já vivemos. A contrapartida que o autor aponta, afinal, é que, mesmo quando trata-se de uma pesquisa quantitativa de investigação de dados, não há como afirmar que não houve subjetividade na investigação.

Assim sendo, recuar às nossas próprias reflexões, como define Cresswell (2002), não é um sinal de falta de rigor metodológico, ou mesmo de ausência de dados, como Stake (2011) e Silverman (2009) reportam quanto à visão positivista de subjetividade. É, ao invés disso, um recurso, uma ferramenta investigativa permitida tão-somente pelo paradigma qualitativo-interpretativista para redimir a anterior e irremissível acusação tipicamente atrelada à pesquisa interpretativa: o excesso de subjetividade. É da natureza da pesquisa qualitativa e da interpretação visar “o aprofundamento da compreensão de um grupo social, de uma organização, etc” (GERHARDT; SILVEIRA, 2009, p.31).

4.4 Triangulação

A triangulação é uma terminologia abrangente dentro do campo da metodologia de pesquisa. Podemos triangular olhares de investigador(es), teorias, metodologias, fontes de dados, dados e, por consequência dos vários tipos existentes, requer uma explicação do porque nos atemos às triangulações de dados e teorias, somente, e não outras. Primeiramente, vamos comentar as vantagens que uma triangulação pode trazer à pesquisa. “A triangulação oferece um meio para que os pesquisadores qualitativos sejam mais discriminadores e criteriosos sobre seus dados” (FIELDING; SCHEREIR, 2001, p.47)³⁸.

Ao triangular investigadores, os dados são investigados, primeiramente, de forma independente uns dos outros. Assim, o mesmo fenômeno é estudado, possivelmente, de diferentes resultados e pontos de vista. A triangulação de dados pode estar ligada à maneira

³⁸ no original - Triangulation offers a mean for qualitative researchers to be more discriminating and discerning about their data

como dados são gerados (ex: métodos como gravações, vídeos, entrevistas, etc.), assim como de um ponto de vista temporal sem alteração de métodos. Nossa visão é unívoca, apesar da participação das orientadoras na pesquisa.

Denzin e Lincoln (2005) propõem que um fenômeno que considera datas – diferenças temporais, e espaços – contextos, gera uma investigação comparativa. Nesse último caso, os dados podem ser coletados ou gerados em diferentes momentos, locais ou sujeitos. A definição anterior está alinhada à natureza de nossa pesquisa, uma vez que utilizamos dados de variados períodos.

Quanto ao aspecto metodológico da triangulação, proposto por Todd em 1979 (GRAY, 2012), consiste-se na combinação de métodos qualitativos e quantitativos para a realização de uma pesquisa. Denzin e Lincoln (2005) a dividem em duas: intramétodo – a utilização de um mesmo método em variados casos – e a de intermétodos – a utilização de diferentes métodos para o mesmo objeto de estudo.

Com base na estrutura da prova e no nosso referencial, partimos para a triangulação de teorias. Para nossa pesquisa, enxergamos a triangulação de dados como a alternativa por “permit[ir] ao pesquisador ter um máximo rendimento teórico usando os mesmos métodos.” (GIBBS, 2009, p.63). Dessa forma, podemos partir de diferentes pontos de vista de um mesmo assunto, sendo eles, o ponto de vista da TSC, da análise do discurso sociocognitiva contextual e das metodologias sociocognitivas de análise de validade de contexto para avaliações a fim de substanciar nossa produção sobre os dados coletados.

4.5 Técnicas

A técnica que utilizaremos para coletar os dados será a leitura, endossada por uma análise de conteúdo temática. Parece tautológico dizer que a leitura será a técnica primária de uma pesquisa, mas dado que existem outras maneiras de conduzir pesquisas que requerem menos leitura e mais atividade documental para gerar dados, por exemplo, tratando-se de uma pesquisa de natureza teórico-exploratória, cremos ser necessário evidenciá-la como técnica.

As dimensões do uso da leitura dependem de sistematicidade procedural com etapas e

funções distintas, como ressalta Salvador (1986). São elas: a leitura de reconhecimento, que serve para localização das obras quanto ao tema; a exploratória, que verifica a pertinência da informação; a seletiva, que objetiva a seleção final do material; a crítica, para a síntese; e a última, a interpretativa, que busca pontos concorrentes entre as fontes selecionadas com os problemas e suas soluções.

Só então podemos iniciar nossa análise temática por meio da transcrição dos documentos relevantes à estrutura da prova (grade de avaliação, modelos de prova e artigos científicos, quando necessário) a fim de investigar (BARDIN, 1979) os elementos sociocognitivos (des)alinhados aos propósitos contextuais da prova. Para tal, três etapas são necessárias: [1] a pré-análise, para obter o material de pesquisa, parcialmente iniciada no referencial teórico; [2] a exploração do material, presente em cada uma das seções necessárias à análise do modelo sociocognitivo de validade de contexto de Weir (2005); [3] tratamento dos resultados, para converter os dados brutos em informações úteis ao nosso quadro analítico, exposto no corpo do texto de análise e nas considerações finais (MINAYO, 2007).

4.6 Componentes Linguísticos de análise

Os componentes linguísticos de nossa análise são os aspectos do insumo das tarefas sob visada sociocognitiva nos âmbitos lexicais, funcionais e estruturais. Utilizaremos dupla abordagem para a análise: a primeira parte do modelo sociocognitivo de análise de contexto elencado em Van Dijk (2012) para a análise linguístico-discursiva dos elementos que compõem a grade de avaliação, suas tarefas e demais documentos; a segunda será o modelo sociocognitivo de análise de validade de contexto de Weir (2005), que não só rege a organização da seção de análise, como também organiza os tópicos ali presentes.

4.7 Procedimentos

Quadro 9 - Procedimentos de pesquisa

Quadro de procedimentos
<ul style="list-style-type: none"> - Investigar um possível construto oral acadêmico - Investigar o exame oral utilizando o modelo de análise de contexto de Van Dijk (2012) junto do modelo sociocognitivo de Weir para validade de contexto (2005) ; - Correlacionar, quando possível, investigar e refletir sobre o construto-alvo do exame oral considerando o contexto simulado e o propósito do exame; - Correlacionar, quando possível, investigar e refletir sobre os descritores presentes na grade de avaliação em relação ao seu construto-alvo, contexto (VAN DIJK, 2012) e alinhamento aos propósitos em termos de validade de contexto (WEIR, 2005) - Apresentar os resultados da investigação acerca do exame oral do IELTS e sua validade de contexto;

Fonte: (WEIR, 2005; VAN DIJK, 2012)

Advertimos os leitores que a ordem da análise liga-se, primariamente, ao modelo de Weir (2005), enquanto o conteúdo da análise apresenta-se de forma polivalente. Isto é, as menções a Van Dijk (2012), e as demais visões de Meurer (2006), Matthiessen e Halliday (2009) e Ramos (2007) aparecem junto das seções de análise elencadas por Weir (2005) uma vez que o instrumento investigado é, primariamente, o exame oral, suas tarefas, grade de avaliação e validade de contexto. Cremos que aplicar o modelo de investigação de contexto de Van Dijk (2012), por exemplo, antes do modelo de Weir (2005) acarretaria em longa exegese que se repetiria tematicamente inúmeras vezes.

4.8 Considerações finais sobre o capítulo metodológico

Encerramos nosso capítulo metodológico afirmando que nossa pesquisa é de natureza qualitativo-interpretativista com base em documentos de variada sorte e que nossa análise se calca em diferentes teorias para investigar e, possivelmente, contribuir com questões referentes à validade de contexto do exame oral do IELTS. A seguir, apresentamos nossa análise.

Capítulo 5 - Uma análise sociocognitiva do exame oral do IELTS

5.1 Análise de potenciais construtos acadêmicos

Um impasse comum na confecção de exames de proficiência para níveis avançados está relacionado à definição de corpus lexical que seja capaz de representar os níveis C1 e C2. Taylor (2011) nos traz vários exemplos referente ao conjunto de testes principais da Cambridge ESOL e, sempre que lida com os referidos níveis, não traz à tona nenhuma bibliografia. Isso não quer dizer, por exemplo, que não haja expectativa alguma quanto ao que se pode definir como um falante avançado. A produção esperada de um candidato à prova oral do *C2 Proficiency* não é menos do que “manter o controle de uma ampla gama de formas gramaticais simples e complexas” (TAYLOR, 2011, p.163)³⁹ O *C2 Proficiency*⁴⁰ é um teste, dentre vários exemplos, que pode ser utilizado para fins acadêmicos, assim como o *C1 Advanced*⁴¹, que, apesar de estar um nível abaixo (C1), também serve para tal propósito.

Em consonância com esses testes, vemos vários editais acadêmicos elencando especificações de níveis de proficiências, além de projetos adequados, para que candidatos de outros países possam ser aceitos em solos anglófonos. O programa desejado, é claro, é o responsável por decidir qual é o nível mínimo de proficiência necessária para que o candidato possa, ou não, participar do programa de graduação, ou pós-graduação, que ele oferece. A maioria dos editais para pós-graduação citam como proficiência mínima o nível B2, e em casos de cursos cuja natureza da informação pode ser cognitivamente mais complexa, são exigidos níveis C1 e, em raros casos, C2.

Sobre a natureza de exames utilizados para fins acadêmicos, Galaczi e French (2011) revelam que o construto que rege as provas relacionadas aos níveis mencionadas no parágrafo é sociocognitivo, para as autoras “falar é visto como tendo uma dimensão cognitiva (ou seja: conhecimento e processamento) e uma dimensão social conceituada em termos de

³⁹ No original: “maintain control of a wide range of simple and complex grammatical forms”

⁴⁰ Conhecido anteriormente como CPE, citado na introdução.

⁴¹ Conhecido anteriormente como Cambridge Advanced English Test of Proficiency.

reciprocidade e co-construção de interação” (2011, p.117)⁴². É notável, a partir da afirmativa das autoras que, mesmo quando o C2 *Proficiency* e o C1 *Advanced* são utilizados para fins acadêmicos, o construto operacionalizado não é uma habilidade acadêmica em si, nem por meio das tarefas, nem por meio dos critérios de avaliação presentes na grade de avaliação desses exames⁴³.

Considerando que o IELTS é utilizado *também* para fins acadêmicos, podemos nos questionar se o construto mobilizado não encontra-se em situação similar à dos testes anteriores. Começamos nossa investigação do construto operacionalizado no exame por meio de uma comparação com o que poderia ser considerado um construto oral voltado para habilidades acadêmicas.

Para essa investigação, utilizamos de diferentes contribuições. A primeira delas é de Rosenfeld, Leung e Oltman (2001), que conduziram uma pesquisa para investigar as demandas linguísticas necessárias para concluir um curso, ou trabalho de curso. Os autores notaram que os alunos mais cognitivamente adaptados às rotinas acadêmicas demonstraram melhor habilidade para lidar com assuntos abstratos, formulação e estruturação de hipóteses.

Em termos de uso da linguagem para resolução de problemas de natureza acadêmica, os autores descobriram que, ainda que a natureza das graduações presentes em seu estudo fosse diversa, as mesmas funções prevaleciam: organizar e elencar informações, opinar e argumentar sobre o que foi dito, descrever objetos, comparar e contrastar, boa pronúncia e precisão para instrutores e também para apresentações. Ainda sobre as atividades comuns ao meio acadêmico, Kim (2006) e Ingram e Bayliss (2007) contribuem para o delineamento do construto oral acadêmico por meio de uma pesquisa voltada à descoberta das três atividades mais comuns em uma sala de aula de nível superior. Suas descobertas apontam para 1) interações em sala de aula envolvendo toda a turma; 2) formulação de perguntas durante a aula; 3) discussões em grupos pequenos. Os entrevistados na pesquisa de Kim (2006) reportaram que a habilidade de conduzir apresentações orais e compreender o que é dito pelos professores no ambiente de sala de aula são, sem dúvidas, as habilidades mais necessárias para o sucesso acadêmico.

⁴² No original: “speaking is seen as having both a cognitive (i.e.: knowledge and processing dimension) and a social dimension conceptualised in terms of reciprocity and co-construction of interaction.”

⁴³ Para mais detalhes ver “Examining Speaking” (TAYLOR, 2011)

Ducasse e Brown (2011) conduziram observações em sala de aula para investigar os tipos de interação caracteristicamente presentes em universidades. Diferentemente dos autores citados nos parágrafos anteriores, as autoras elencaram dezesseis tipos de interação que consideravam inerentes ao discurso e prática acadêmica e compararam-nas às funções equivalentes no exame oral do IELTS. As autoras constataram que, das dezesseis funções investigadas, apenas sete apareceram nos exames orais do IELTS investigados.

Quadro 10 - Relação linguagem acadêmica e emergência no IELTS

Tipo de função	Função	Frequência	No IELTS
Informação 1	Prover informações, ideias ou opiniões	5	x
Informação 2	Defender as próprias ideias ou opiniões	5	x
Informação 3	Elaborar as próprias ideias ou opiniões	1	
Informação 4	Sugerir	4	
Interação 1	Desafiar ideias	3	
Interação 2	Justificar/prover suporte a outro	2	
Interação 3	Concordar/Discordar	4	x
Interação 4	Quantificar/modificar	1	x
Interação 5	Persuadir	2	
Interação 6	Pedir informação, ideias ou opiniões	4	x
Interação 7	Solicitar elaboração/justificativas	1	x
Interação 8	Desenvolver/modificar a ideia de outrém	3	
Interação 9	Negociar significado	5	x
Gerenciamento 1	Iniciar interação	1	
Gerenciamento 2	Mudar de assunto	1	
Gerenciamento 3	Concluir um argumento/decidir	2	

Fonte: (DUCASSE; BROWN, 2011, p.12)

Kettle e May (2012) catalogaram palestras/aulas, tutoriais e linhas guias de avaliação para a realização de tarefas com a colaboração de alunos do primeiro ano de graduação em nível superior e constaram que a maioria das habilidades utilizadas foram para responder perguntas do palestrante/professor e pedir por esclarecimento e/ou confirmação, questionando os tutoriais dados. Os colaboradores da pesquisa de Kettle e May (2012) não apresentaram nenhum comportamento específico, em termos interacionais e/ou linguísticos quando se tratando das linhas guias da tarefa, diferentemente do que fizeram com palestrantes ou

professores, mas mostraram-se engajados nas discussões para co-construir a tarefa descrita nas linhas guias de avaliação.

Uma vez que é ponto pacífico tanto para a avaliação sociocognitiva quanto para a linguística sociocognitiva que um construto em sua forma operacional não representa a totalidade de seu construto mental, há de se interpretar e contrastar o uso da língua em contexto de avaliação e fora de contexto de avaliação (ex: interações reais). Brooks e Swain (2014) compararam o desempenho oral de candidatos do TOEFL durante a prova ao seu desempenho em situações reais envolvendo apresentações orais, discussões e conversas em grupos pequenos dentro e fora de sala de aula. O construto operacionalizado no exame oral foi mais complexo tanto em termos sintáticos (variabilidade das estruturas) quanto precisos (número e características dos erros cometidos). Tal pesquisa também demonstrou que houve variação de registro dentro das tarefas do teste, prevalecendo uma amostra formal, com nenhuma incidência de perguntas, em comparação às feitas em sala de aula. O último tópico investigativo de Brooks e Swain (2014) revela-se ligeiramente tautológico, uma vez que o TOEFL é estritamente monológico, mas a observação não deixa de ter valor por ser um ponto de reflexão sobre o que podemos ou não incorporar em um teste de proficiência para potencializar a operacionalização de um construto de maneira mais alinhada às práticas sociais de seu contexto original. Afinal, perguntas são uma notável parte das interações de cunho acadêmico.

Podemos, então, com base no exposto e nos apontamentos de Inoue, Nakatsuhara, Lam e Taylor (CRELLA, 2018) resumir as habilidades mais relevantes ao ambiente acadêmico como:

- organização de conteúdo/argumento
- precisão na produção (clareza, recepção, etc.)
- a habilidade de falar por longos períodos para um determinado grupo de pessoas
- domínio de variadas habilidades interacionais: perguntar e responder perguntas; dar ou substanciar opiniões; concordar e discordar; descrever, comparar e listar/resumir; falar por longos períodos de tempo;
- habilidade para lidar com vários contextos discursivos: entrevistas ou instruções face-a-face (mediador da informação e aprendiz), discussões em grupos pequenos, discussões com base em aulas e/ou conteúdo especificado, apresentações formais e informais (INOUE et al, 2018, p.12)

No entanto, definir as habilidades mais utilizadas em contexto acadêmico não é o mesmo que definir as demandas que subjazem atividades acadêmicas num nível sociocognitivo. De acordo com Taylor (2011), conteúdos programáticos para cada nível, salvo para os níveis C1/C2 podem ser encontrados em textos como QECR (COUNCIL OF EUROPE, 2001), Waystage (1998), Vantage (2001) e similares, em que itens lexicais, expectativa de produção oral/escrita e funções da linguagem comumente esperadas, inclusive em nível de dificuldade, são listadas. No entanto, quando tratamos de um construto oral de inglês acadêmico, o material é relativamente esparsos. Inoue, Nakatsuhara, Lam e Taylor (CRELLA, 2018) apontam que as demandas acadêmicas podem ser inferidas a partir de três diferentes fontes: habilidades de estudo e recursos (eletrônicos, ou analógicos) disponíveis, ementas de cursos de inglês para fins acadêmicos e documentos publicados voltados para a investigação de letramento acadêmico.

A investigação desses autores aponta para os contextos discursivos que mais comumente mobilizam as funções, interações e gerenciamento de interações pertinentes ao contexto acadêmico, ou seja, segundo Van Dijk (2012), isso constitui um modelo mental do contexto subjacente à atividade da fala. Bygate (1987) configura tal modelo mental por outra nomenclatura, chamando-a de rotina, mas que, em essência, se encontra em consonância com o conceito de modelo mental de contexto de Van Dijk. A rotina, tal como o modelo mental, é o esquema mental cognitivo que se acopla ao construto de habilidade oral para “afunilar” o teor sociocognitivo de uma determinada interação.

Inoue et al (2018) listam os seguintes contextos discursivos como exemplos arquetípicos de interações acadêmicas. Os autores concebem que o gênero (como contexto [VAN DIJK, 2012]) apresentação possui características hierárquicas mais marcadas e menos interação possível, enquanto o seminário tem natureza mais cooperativa, colaborativa e simétrica.

Ao adentrar a investigação das rotinas sociocognitivas inerentes à formação dos contextos academicamente discursivos, os autores apontam para seis características e/ou rotinas⁴⁴.

⁴⁴ A variabilidade das rotinas mostrou-se baixa, nos estudos dos autores, uma questão de estilo e hábito. Em essência, as funções subjacentes às atividades a serem desempenhadas eram as mesmas. Em outras palavras, para nós, todos os colaboradores investigados, de uma forma ou de outra, passaram por processos cognitivos similares e por isso, optamos por resumir o quadro descrito na pesquisa de Inoue et al (2018).

Quadro 11- Características de apresentações e seminários

Apresentações	Discussão em seminário
<ul style="list-style-type: none"> ● Individual ou em grupo ● Apresentação sobre um tópico ● Apresentação sobre um artigo ● Apresentação de um projeto de pesquisa desenvolvido pelo aluno 	<ul style="list-style-type: none"> ● Grupos pequenos vs. toda a turma ● Discussão sobre uma leitura ● Discussão sobre um conteúdo da aula ● Tarefas colaborativas

Fonte: (INOUE et al, 2018)

Para as apresentações, o primeiro é a organização/estruturação da apresentação com ou sem o auxílio de notas; o segundo é o uso da voz com ou sem variação tonal, controle da velocidade da fala (taxa ilocucionária), variação de tom com ou sem pausas, variação de volume, entonação e presença ou ausência de ênfase; o terceiro é referente ao léxico utilizado em apresentações, que difere da linguagem corriqueira comumente utilizada pelos interactantes; o quarto é a presença ou ausência de linguagem corporal, incluindo a leitura de anotações, uso de gestos e contato visual com os espectadores; o quinto é o uso de algum aporte visual para auxiliar a compreensão de um assunto; e último é o foco da apresentação, que pode ser interativo ao fim, em que o apresentador responderá perguntas, ou expositivo, sem interação com os espectadores.

O estudo de Inoue et al (2018) revela muitos pontos nevrálgicos característicos da interação acadêmica, além de abordar os princípios sociocognitivos, isto é, os processos socioculturais que edificam a rotina de apresentação e sua formação discursiva. Dos itens descritos, o que mais chamou nossa atenção foi a maneira como os autores organizaram as interações de acordo com o uso característico da língua em contexto acadêmico. Há claros pontos de cisão entre uma interação corriqueira, como a exposição de um fato, e uma apresentação. A segunda, por exemplo, preza pela organização prévia da informação, que não tende a ocorrer na fala espontânea, a preocupação e controle em relação à taxa ilocucionária a fim de garantir boa compreensão do tópico, as palavras, frases e linguagem corporal tipicamente utilizadas em apresentações (incluindo a associação de ambas, como o ato de apontar para uma frase em um eslaide), o uso de recursos visuais para acompanhar o ato da fala e, por fim, mas não tão incomum, a disponibilidade para lidar com seus ouvintes fora de um contexto de tutelação.

Quanto às rotinas discursivas presentes em seminários, Inoue et al (2018) também elencam seis características particulares. Diferentemente da apresentação, o seminário pressupõe maior possibilidade de interações. Portanto, novas rotinas são inerentes a esse contexto interacional. Os autores elencam como primeiro ponto característico a necessidade de expressar ideias, apresentar argumentos citando fontes, fala clara, assim como comentários sobre as leituras feitas. O segundo é referente à natureza interacional de um seminário, cuja presença de perguntas é tolerada e aceitável, tanto entre alunos e apresentadores, como entre alunos e professores-mediadores. A terceira é a possibilidade de interação bilateral entre apresentador e plateia durante a apresentação de seu discurso, em que também destaca-se a responsabilidade que os ouvintes têm em acompanhar o que é apresentado, assim como responder perguntas de seus professores e outros colegas (a depender dos comentários feitos durante a interação). A quarta característica é referente à possibilidade de concordar ou discordar com o que foi posto na apresentação que pode ou não envolver atos de persuasão da parte dos interactantes; ademais, relaciona-se também a esse tópico, da parte da plateia, possíveis questionamentos quanto à relevância das ideias promovidas e concatenadas pelos apresentadores, assim como pedidos de esclarecimento e contribuições à apresentação, quando cabível. A quinta característica marcante relaciona-se à um aspecto sociocognitivo particular, que é a noção de autoria e senso de contribuição, assim como a possibilidade de interrupção e/ou retificação. A sexta característica discursiva de um seminário em sala de aula é o marcante uso do léxico referente ao tópico e a presença de um resumo geral para auxiliar os discentes envolvidos a assimilar o que foi dito.

Quadro 12: Resumo dos elementos inerentes à apresentação e ao seminário

<p>Interação com uma audiência (principalmente unidirecional)</p> <ul style="list-style-type: none"> ● Organização de conteúdo/estrutura ● Clareza na apresentação ● Consciência do discurso ● Controle paralinguístico ● Uso de material visual ● Interação funcional 	<p>Interação em um grupo (principalmente multidirecional)</p> <ul style="list-style-type: none"> ● Planejamento e gestão do conteúdo ● Interação funcional (mas mais rica e variada)
--	--

Fonte: (INOUE et al, 2018)

Com base no que foi exposto, é possível delinear o que seria crucial, de acordo com Inoue et al (2018) e Ducasse e Brown (2011), a um construto acadêmico:

Quadro 13: Elementos cognitivos operacionalizados

Elementos interacionais constituintes do construto oral acadêmico segundo Inoue et al (CRELLA, 2018)	Elementos interacionais comumente utilizados em exames orais de proficiência
<p>Pesquisa das universidades</p> <ul style="list-style-type: none"> ● organização de conteúdo/argumento ● clareza na produção ● habilidade de produzir turnos longos ● domínio sobre várias habilidades interacionais: <ul style="list-style-type: none"> - perguntar/responder perguntas - dar/defender opinião - concordar/discordar - descrever/comparar/resumir ● habilidade de lidar com uma variedade de contextos discursivos, principalmente: discussões em grupo e apresentações formais/informais 	<p>Pesquisa de literatura</p> <ul style="list-style-type: none"> ● organização de conteúdo/argumento ● clareza na apresentação (clara, precisa) ● habilidade de produzir turnos longos ● domínio sobre várias habilidades interacionais: <ul style="list-style-type: none"> - perguntar/responder perguntas - dar/defender opinião - concordar/discordar - descrever/comparar/resumir ● habilidade de lidar com uma variedade de contextos discursivos: entrevista ou conversa 1 a 1, discussão em grupos pequenos, discussão com a turma, apresentação formal/informal

Fonte: Fonte: (INOUE et al, 2018)

Com base na definição do que é necessário para que se operacionalize um discurso acadêmico autêntico em relação ao seu contexto, investigamos a estrutura do exame oral do IELTS a fim de compreender a relação entre as tarefas, sua grade de avaliação e os elementos

que compõem, sob lente sociocognitiva, a validade de contexto oral de um exame de proficiência acadêmico.

5.2 Análise das instruções, comandos, enunciados e formato das respostas

Para investigar o formato das respostas do exame oral do IELTS, devemos primeiro isolá-las, identificar as instruções, enunciados e comandos para, então, investigá-las à luz de nosso referencial. Tratando do assunto em questão, Weir (2005) defende que a ordem dos itens é crucial para que um teste operacionalize o construto-alvo de maneira alinhada ao seu propósito. Taylor (2011) pondera que testes de proficiência geral, cujas tarefas e grade de avaliação procuram avaliar o desempenho de vários níveis diferentes de proficiência, devem organizar a ordem das questões de tal forma que a dificuldade não seja sentida de maneira abrupta, mas gradativa. A autora defende que a distribuição de funções, a quantidade de amostra produzida, as oportunidades de gerenciar o diálogo dadas aos candidatos e a distribuição de direitos e deveres conversacionais são diretamente afetadas pela tarefa. Ademais, também atribui-se ao formato da resposta, íntima ligação com o propósito geral do teste e, em decorrência disso, potenciais entraves relacionados a diferentes formatos de resposta devem ser considerados à luz do propósito do teste.

Bygate (2001) relaciona o alinhamento da tarefa ao propósito com uma potencial fonte de injustiça, uma vez que podem haver erros de pontuação decorrentes do efeito causado pelo enunciado e dificuldade da tarefa ao desempenho de um candidato. Para lidar com tal problema, Alderson, Clapham e Wall (1995) defendem que a adoção de diferentes métodos de geração de dados pode reduzir vieses (como maior familiaridade com uma determinada função discursiva) e incentivar o uso de variadas estratégias cognitivas de comunicação, o que pode resultar num leque maior de amostras e, conseqüentemente, melhores inferências sobre o desempenho do candidato avaliado. Segundo Galaczi e French (2011), além de uma maior amostra, múltiplas possibilidades de produção advindas de diferentes formatos de resposta e temas podem reduzir a variância inerentemente ligada à competência comunicativa de um candidato, o que seria inviável, caso somente um método estivesse sendo utilizado.

Weir (2005) aponta que, além dos aspectos citados, que são relevantes ao formato da resposta, o que também se deve considerar é se a adequação e interação envolvidas no formato da resposta são apropriadas para o nível de proficiência que vai ser avaliado. Tal reflexão de Weir é relevante para a confecção de um exame, pois diferentes tipos de formato de resposta e interação geram diferentes cargas cognitivas. Sendo assim, uma distribuição dessa carga pode vir a ser útil em níveis de proficiência cuja habilidade do candidato sendo testada seja muito limitada. Há várias maneiras de se distribuir a carga cognitiva de uma prova ao longo de suas tarefas. A primeira, que está ligada à investigação do formato da resposta, é o tipo de interação presente na tarefa.

Há dois tipos de interação: a monológica e a dialógica. A monológica define-se por apresentações orais (como em turnos longos de fala) e/ou respostas a computadores e/ou gravações. A dialógica caracteriza-se como uma espécie de entrevista, dramatização interativa, interação candidato-mediador para uma discussão, candidato-candidato, ou interação entre mediador-candidato para o mesmo fim. Entender os tipos de interações em um exame nos ajuda a investigar como as oportunidades interacionais de uma tarefa são simuladas e de que forma os processos cognitivos mobilizados relacionam-se com o construto que a prova deseja evocar e, conseqüentemente, o contexto simulado. A interação candidato-candidato, por exemplo, tende a ser mais simétrica, mobiliza funções da fala com mais facilidade e incentiva a troca de turnos (LAZARATON; FRANTZ, 1997).

A interação candidato-mediador mostra-se, em contrapartida, mais assimétrica em termos de dominância (quem controla a interação), contingência (a liberdade para desenvolver os tópicos livremente) e orientação de objetivos (tarefa essa incumbida somente ao mediador) (YOUNG; MILANOVIC, 1992). Nesse sentido, French (2003) argumenta que interações candidato-candidato permitem mais oportunidades de amostra e, corroborando a pesquisa de French (2003), Brooks (2009) aponta, também, para a gama de interações complexas advindas de tal dinâmica, como mais elaboração e busca por esclarecimento, mais sentenças finalizadas, mais referências aos participantes da conversa (caso haja mais do que dois) e maior presença de paráfrases, enquanto a examinador-candidato pode limitar a troca de papéis e a fuga do assunto determinado pelo examinador. Todavia, de acordo com Weir (2005), uma interação sem reciprocidade prejudica a validade de um teste.

Dessa forma, para garantir uma produção que possua mais caráter sociocognitivo em um exame oral de proficiência, é necessário a implementação de interações que permitam a co-construção de diálogos. Tal como Van Dijk (2012) assevera, o discurso é dinâmico e, conseqüentemente, uma interação demasiadamente controlada que não permita acesso ao dinamismo característico da fala a aflorar, pode prejudicar o desempenho dos candidatos.

Pensando nos dois tipos de interação descritos anteriormente, conseguimos traçar dois padrões de interação de acordo com as entrevistas transcritas em Seedhouse et al (2004). O padrão dialógico nas partes 1 e 3, e o padrão monológico na parte 2. Cada um dos formatos presentes mobiliza diferentes processos cognitivos, a serem desenvolvidos mais a fundo na seção de modo de discurso e aspectos funcionais do exame.

A parte 1 caracteriza-se por tarefas cujo caráter é, majoritariamente, informacional. Segundo as categorias de Bygate (1987), o caráter informacional é aquele que lida com informações pessoais, descrições e/ou breves opiniões. A organização das tarefas se dá a partir do preâmbulo apresentado pelo mediador ao início do exame, que contém as instruções iniciais para essa parte “Eu gostaria de te fazer algumas perguntas sobre você” (SEEDHOUSE; EGGBERT, 2004, p.40)⁴⁵. O enunciado e os comandos são apresentados junto do tema das perguntas, tal como apresentado no capítulo metodológico. Uma análise do léxico que compõem as tarefas encontra-se na seção 5.10, tanto para a parte 1, como para as partes 2 e 3. A dificuldade no exame oral do IELTS apresenta-se de forma progressiva, pois o conteúdo apresentado nas tarefas relaciona-se, primariamente, a informações pessoais e, aos poucos, afasta-se desse nível do discurso para tarefas que mobilizam opiniões sobre um determinado assunto dentro do que o QEQR (COUNCIL OF EUROPE, 2001) considera como conhecimento abstrato.

Outros dois conceitos de Bygate (1987) que podem ser úteis à análise dos formatos das respostas são o interacional e o interacional-gerencial. O segundo, além do informacional descrito no parágrafo anterior, é interacional. Logo, tal conceito lida com funções da língua ligeiramente mais complexas, tais como: persuadir, concordar ou discordar junto de sustentação argumentativa durante a interação. Por fim, a última função, que é a interacional-gerencial, relaciona-se às responsabilidades dialogais, tal como iniciar, ou terminar, uma conversa, mudar de assunto, dentre outras. Uma análise mais aprofundada das

⁴⁵ no original - I'd like to ask you some questions about yourself.

características das tarefas e partes do exame encontra-se na seção de aspectos funcionais, nos quadros 31, 32, 33 e 34, e classifica as amostras de acordo com as categorias de Bygate.

Galaczi e French (2011) constataram que interações de turnos de fala longo e interações em pares (face-a-face, individual) tendem a mobilizar mais funções informacionais, enquanto discussões abertas tornam o diálogo mais livre, e são capazes de mobilizar as três funções da conversa-avaliativa descritas por Bygate (1987). Tal prática possibilita mais amostras para análise e amplia as possibilidades de obtenção de evidências para o avaliador. Parece-nos, contudo, que discussões abertas ampliam a validade, mas diminuem a praticidade. Em um exame de larga escala, a praticidade acaba sendo privilegiada (FULCHER; DAVISON, 2007).

Lazaraton e Frantz (1997), contudo, demonstram que não são todos os tipos de interação face-a-face que potencializam o aparecimento das funções elencadas por Bygate (1987), principalmente no caso daquelas cujo mediador-avaliador apresenta menos funções mediadoras. Isso parece indicar que o formato de respostas empregado no exame oral do IELTS tem potencial para limitar as funções do usuário e tornar as interações mais controladas. A vantagem disso é uma análise mais rápida, com assuntos que elencam funções específicas de maneira pontual, mas que deixam de investigar aspectos intrinsecamente relacionados à validade de contexto acadêmica, assunto a ser desenvolvido adiante.

Na parte 2 podemos notar que a interação entre o mediador e o candidato da prova muda. Nessa parte, a instrução é dada oralmente

Agora, vamos para a parte seguinte. Eu vou te dar um assunto e gostaria que você falasse sobre isso por um ou dois minutos. Antes de falar, você terá um minuto para pensar sobre o que vai dizer e pode fazer algumas anotações, se desejar. (SEEDHOUSE; EGGBERT, 2004, p.41)⁴⁶.

⁴⁶ no original - Now, let's move on to the next part. I'm going to give you a topic and I'd like you to talk about it for one or two minutes. Before you talk, you'll have one minute to think about what you're going to say, and you can make some notes if you wish.

A instrução da tarefa é atribuída, e o enunciado, “Eu gostaria que você descrevesse uma viagem que você fez uma vez” (SEEDHOUSE; EGGBERT, p.42)⁴⁷, atribuído oralmente, é seguido dos comandos no cartão do candidato

Cartão de tarefas

Eu gostaria que você descrevesse uma viagem que você fez uma vez.

- 1 - Para onde você foi
- 2 - Por que você foi lá
- 3 - Como você viajou
- 4 - O que você mais gostou
- 5 - Se você repetiria esta viagem e por quê
- 6 - Você geralmente gosta de viajar? (SEEDHOUSE; EGGBERT, p.42)⁴⁸

Dessa forma, o candidato deve discorrer sobre uma experiência junto das funções discursivas que gradativamente se apresentam ao longo da tarefa. A primeira questão da tarefa tem caráter informacional, a segunda tem caráter informacional, a terceira tem caráter informacional, a quarta tem caráter opinativo-informacional, enquanto a quinta tem caráter especulativo. A pergunta final feita pelo examinador apresentada ao fim da parte 2 demonstra-se, primordialmente, informacional, uma vez que pede uma opinião, somente.

Na parte 3 há um misto de tópicos de com diversos tipos de funções discursivas complexas. As instruções são dadas pelo mediador-avaliador da seguinte maneira: “[Eu] gostaria de discutir com você uma ou duas questões gerais relacionadas a isso.” (SEEDHOUSE; EGGBERT, p.42)⁴⁹. Os enunciados e comandos são dados de forma conjugada, tal como na parte 1 - a temática, dada pelo enunciado, é seguida dos comandos. A discussão parece levar o candidato a divagar, especular, descrever e julgar os tópicos propostos, enquanto o examinador age somente em função da manutenção das interações que ali ocorrem. Tal prática configura-se como característica sociocognitiva do exame oral por permitir que a interação se construa por meio da língua, na troca de experiências entre os candidatos e seus mediadores. Para Skehan (2001) e Galaczi e French (2011), a variação de papéis de cunho interacional e funcional licencia maior emergência de funções discursivas da

⁴⁷ no original - I'd like you to describe a trip you once went on

⁴⁸ O que foi escrito no cartão de tarefas não é apresentado explicitamente no texto de Seedhouse e Eggbert (2004). Inferimos as perguntas a partir da amostra que os candidatos apresentaram em sua transcrição.

⁴⁹ no original - I'd like to discuss with you one or two more questions related to this

linguagem, bem como assegura amostras mais autênticas e menos assimetria entre os participantes. As figuras 10, 11 e 12 organizam tais categorias e encontram-se na seção de análise funcional do insumo do exame.

Bachman (1990) e Bachman e Palmer (1996) destacam que é a partir das interações suscitadas nos exames orais que somos capazes de fazer inferências sobre as habilidades comunicativas dos candidatos além da situação interacional simulada no exame. Um teste cuja interação é somente monológica deixa de gerar oportunidades de operacionalização de habilidades sociocognitivas relevantes para que mais facetas dos construtos-alvo se revelem. As instâncias da interação são intrinsecamente ligadas a esquemas mentais (VAN DIJK, 2012) que acompanham um determinado construto-alvo e, quanto mais autêntica for a mobilização do construto, mais esquemas podem vir à tona e, conseqüentemente, mais válida a amostra se torna. No entanto, há mais coisas a se considerar além das inferências, um exemplo está nos trabalhos de Fulcher e Davidson (2007), em que claros alertas sobre as potenciais conseqüências desastrosas, tal como Weir (2005) e Bachman (1990), de inferências interpretadas a partir de teorias de proficiência gerais para fins específicos, além de seus efeitos negativos na validade e confiabilidade de um teste.

Em consonância com as afirmativas do parágrafo anterior, porque o teste utiliza interações face-a-face, há movimentação de recursos cognitivos de cunho interacional e, portanto, maiores chances de discursos autenticamente co-construídos.

As três partes do exame oral não tem alterações temáticas nas versões transcritas por Seedhouse et al (2004), mas White (2020) adverte que uma gama de assuntos permeia a temática da prova. Uma vez que não investigaremos um livro didático no corpo de nossos documentos, partimos dos pressupostos teóricos de Luoma (2004) para nossa análise do exame, os quais elucidam uma questão presente em vários testes de proficiência: um teste não deve se apegar ao conteúdo das perguntas, salvo em casos voltados para fins extremamente específicos, como aviação, mas sim, às funções nelas elencadas.

Quadro 14 - Resumo do formato das respostas: método de geração de dados e propósito das tarefas

	Tipo de interação	Função(ões) predominante(s)	Tipo de interação e formato da resposta	Propósito da tarefa
Parte 1	dialógica entre candidato e mediador	informacional descritiva elaborativa emissão de opinião	dialógica direta	Esta parte do teste se concentra na capacidade de comunicar opiniões e informações sobre tópicos do dia-a-dia e experiências ou situações comuns, respondendo a uma série de perguntas.
Parte 2	monológica com candidato produzindo turno de fala longo	informacional descritiva elaborativa emissão de opinião especulativa	monológica direta ⁵⁰	Esta parte do teste enfoca a habilidade de falar longamente sobre um determinado tópico (sem instruções adicionais do examinador), usando linguagem apropriada e organizando ideias de forma coerente. É provável que os participantes do teste precisem se basear em sua própria experiência para concluir a curva longa.
Parte 3	dialógica entre candidato e mediador	elaborativa emissão de opinião especulativa comparativa	dialógica direta	Esta parte do teste se concentra na capacidade de expressar e justificar opiniões e de analisar, discutir e especular sobre questões.

Fonte: (BYGATE, 1987; LUOMA, 2004; TAYLOR, 2011; IELTS, 2021)

Um aspecto digno de nota sobre a estrutura do exame oral do IELTS é que, segundo Taylor (2011), diferentes métodos são maneiras de se obter maior gama de amostras. Porém, de acordo com o breve desenho do que viria a ser um construto acadêmico, não se pode dizer que as interações suscitadas em seu exame oral são suficientes para garantir uma amostra de construto suficientemente generalizável para aferir o construto subjacente à prática social acadêmica, bem como seus construtos participantes.

⁵⁰ Há de se considerar que, apesar de se tratar de uma interação face a face, o formato da parte 2 muito se assemelha ao de um exame semi-direto.

Em contrapartida, consideramos a ausência de interações mais apropriadas para o contexto acadêmico relativamente compreensível. O IELTS se propõe a certificar candidatos de nível mínimo B1 a C2. Logo, há de se levar em conta, primeiramente, a praticidade temática das questões da prova, especificamente porque o IELTS trata de um público majoritariamente heterogêneo. Partindo desse pressuposto, podemos inferir que os efeitos que tarefas academicamente mais contextuais trariam ao contexto da prova acarretariam num aumento de demanda e carga cognitiva que candidatos de nível B1, por exemplo, seriam capazes de lidar. Field (2011) adverte que suscitar diferentes funções e interações em um teste pode ter resultados positivos, mas também pode acabar por limitar as possibilidades de interação de um candidato, gerando, assim, uma espécie de viés. No exemplo descrito, o viés favorece àqueles que, como Van Dijk (2012) explica, já possuem arcabouço esquemático mental, ou seja, o construto mental, para lidar com as demandas cognitivas provocadas nas instâncias de interação do exame.

Creemos que a grande heterogeneidade do perfil médio dos candidatos influencia o formato e ordem das questões. Posto isso, tarefas cujo propósito é acadêmico, ou profissional, mas que partem de um construto de proficiência geral, podem ter a validade de contexto fragilizada. Desenvolveremos mais sobre o perfil dos candidatos nas seções adiante.

5.3 Análise de propósito das tarefas

Como já discutido no referencial teórico, o propósito de uma tarefa deve estar relacionado ao propósito geral do exame (WEIR, 2005). As tarefas devem refletir instâncias sociais, representando interações situadas em um nível micro, ou seja, um fragmento discursivo (VAN DIJK, 2012) que aluda ao nível macrosociológico por meio da ativação de esquemas mentais e construtos participantes paralelos ao construto-alvo da avaliação.

As tarefas, então, têm a obrigação de mobilizar o que o candidato deve ser capaz de fazer com a língua-alvo. Se essa premissa for respeitada, define-se o propósito da tarefa (LUOMA, 2004). Creemos que seja evidente a relação existente entre validade de contexto e cognição em razão de que, o primeiro, é parte do construto cognitivo de um ponto de vista da formação da cognição social, pois resulta de informações culturais e situacionais sobre as

linguagens e interações humanas. Consequentemente, a validade cognitiva, ou seja, os processos cognitivos mobilizados durante a realização de um exame, é afetada, caso o contexto simulado na prova não mobilize os processos cognitivos socialmente alinhados aos contextos originais para operacionalizar o construto que se deseja investigar. Para assegurar que a validade de contexto seja refletida também na validade cognitiva, e vice-versa, a implementação de múltiplas instâncias práticas dentro das tarefas de um exame pode auxiliar na mobilização de diferentes esquemas mentais de micro e macroprocessamento da linguagem. O microprocessamento é a relação simbiótica existente entre as escolhas que fazemos em relação ao propósito e o processamento mental necessário para realizar uma determinada tarefa, enquanto o macro alude ao processamento do contexto de ocorrência (WEIR, 2005).

Um propósito claro facilita a percepção dos objetivos de uma tarefa e também do monitoramento pessoal, ou auto-monitoramento do candidato (KRASHEN; TERRELL, 1980), que é relativo às estratégias cognitivas relevantes para uma produção alinhada a possíveis demandas linguísticas, interacionais e funcionais. Um propósito que desvia o candidato do objetivo de uma tarefa não só pode prejudicá-lo em termos de planejamento, mas também pode resultar na piora de seu desempenho durante o exame. (TAYLOR, 2011).

O propósito é expresso pelas instruções, enunciados e comandos de uma tarefa, assim como pela escolha das palavras dos três componentes da tarefa ao nível de proficiência dos candidatos ao exame. Esse aspecto é particularmente chamativo no exame oral do IELTS por alguns motivos. Primeiramente, como já mencionado, o exame certifica níveis B1, B2, C1 e C2. A variação existente entre esses níveis, em termos cognitivos, é díspar, pois, ainda que o candidato tenha a capacidade de processar informações abstratas e não familiares a sua primeira língua, tal processamento não irá necessariamente ser expresso na língua-alvo, possivelmente em decorrência da falta de léxico e/ou falta de familiaridade com cosmologia sociocognitivo-cultural da língua-alvo. O construto a ser manifesto tornar-se-ia imprevisível (WEIR, 2005). Para lidar com esse problema, Galaczi e French (2011) apontam como prática comum dos elaboradores, a simplificação de instruções, enunciados e comandos, juntamente da inserção de elementos interacionais, funcionais, ou ambos, que intensifiquem, mitiguem ou mesquem usos da linguagem de caráter simples e complexo.

Tal como Luoma (2004) defende, simplificar instruções, enunciados e comandos primando pelas funções interacionais e cognitivas da linguagem pode favorecer candidatos mais hábeis, dado que esses possuem mais percepção quanto às oportunidades geradas pelas “palavras-chave” sem desfavorecer candidatos de baixa proficiência, que apresentariam dificuldades em manter o diálogo e poderiam se sentir desencorajados a se engajar com a tarefa pela complexa natureza do teste, caso não simplificado fosse. Dessa forma, pode-se evidenciar as habilidades cognitivas do candidato por meio do formato, funções, interações, demandas e ordem das questões, ou seja, o contexto experiencial já presente em suas mentes.

Para esclarecer um pouco mais sobre funções simples e complexas, discorreremos brevemente sobre alguns autores a fim de retomar o referencial teórico e direcionar os recursos a serem elencados neste parágrafo para a investigação das tarefas presentes nas amostras investigadas. Brown e Yule (1983) organizam funções por ordem de complexidade, uma das ferramentas que poderia ser implementada em tarefas cujo teste avalia múltiplos níveis. Em ordem de complexidade, os autores organizam as funções *descrever, instruir, narrar um evento, expressar e/ou justificar opiniões*. Similarmente, Bygate (1987), já anteriormente citado, organiza, de maneira temática e gradativa, duas categorias para guiar diálogos: a primeira é a conversa-factual (*factually-oriented talk*), que compreende *descrição, narração, instrução e comparação*; a segunda é a conversa-avaliativa (*evaluative-talk*) que compreende *explicação, justificativa, produção e decisão*.

Outrossim, ainda que seja possível reduzir a complexidade lexical de um exame e assegurar a presença de funções complexas subjacentes às suas tarefas, outro problema revela-se, este sendo nosso principal foco de pesquisa. As modalidades acadêmica e geral do exame oral do IELTS são utilizadas para diferentes propósitos, mas avalia seus candidatos por meio de um mesmo modelo de exame oral. A incongruência, destarte, dá-se na observável generalização do construto do exame oral em relação à validade de contextos contemplados. Podemos nos perguntar se o construto-alvo não seria, por exemplo, em vez de inglês para fins acadêmicos, inglês geral. O mesmo questionamento se aplica para a modalidade geral da prova.

A definição de contexto como manifestação discursivo-social de Van Dijk (2012), que envolve modelo mental, nos conduz a perceber o problema epistemológico presente nas avaliações de construto mediadas pela prova oral do IELTS, já que o instrumento utilizado

não suscita as interações mais relevantes ao contexto que se propõe a avaliar. Contextos são inerentes a gêneros discursivos e, tal como Vygotsky (1986) postula, discursos são moldados e remoldados em suas particulares instâncias interacionais. Isto é, os processos sociocognitivos particulares pertencentes a um determinado contexto refletem-se em esferas sociais, meios de comunicação, hierarquias e aspectos culturais, impactando diretamente as escolhas lexicais, funções discursivas características, objetivos interacionais, dentre outras infinitas possibilidades de situações sociais, que demandam aspectos linguísticos e discursivos peculiares. A organização do propósito de uma tarefa deve alinhar-se ao propósito do teste e ao objetivo final do candidato.

Retomando a breve investigação feita sobre o que seria um construto oral acadêmico à luz, principalmente de Inoue et al (2018), podemos notar que nem todas as características tidas como pertencentes ao discurso acadêmico são contempladas nas tarefas do exame oral do IELTS. A investigação do construto referente ao inglês geral não se faz necessária nesse caso porque a prova já se mostra parcialmente alinhada às interações características de um de seus propósitos. Logo, por se tratar do mesmo exame, pelo menos quanto ao construto oral, pode-se inferir, a partir da análise de propósito das tarefas, que a validade de contexto do exame oral do IELTS encontra-se prejudicada.

O quadro 15 aborda as características sociocognitivas relevantes ao construto oral acadêmico. Com base nas perguntas transcritas no capítulo metodológico, a partir dos artigos de Seedhouse et al (2004), analisamos os tópicos elencados por Inoue et al (2018) e os contrastamos por meio de uma análise sociocognitiva e funcional. Identificamos a função de cada uma das tarefas com base em Luoma (2004), Bygate (1987) e Brown e Yule (1987) para ter melhor compreensão das tarefas do exame, de seus modelos funcionais e, conseqüentemente, da sua validade de contexto.

Quadro 15 - resumo do construto acadêmico e da análise de validade de contexto das tarefas do IELTS

CONSTRUTO ACADÊMICO X TAREFAS DO IELTS	
FUNÇÕES, CONSTRUTOS E INTERAÇÕES	CONTROLE (PRESENTE/AUSENTE)
Organização de argumentos	PRESENTE
Clareza na transmissão	PRESENTE
Habilidade de fala por longos períodos	PRESENTE
Leque de habilidades interacionais	PARCIALMENTE PRESENTE
Contextos discursivos variados	AUSENTE

Fonte: (INOUE et al, 2018)

5.4 Análise da distribuição do peso nas tarefas

Quando discutimos pesos, estamos lidando com uma maneira de atribuir características mensuráveis a cada tarefa da prova. O peso, em outras palavras, expressa a relação de todos os descritores de uma grade de avaliação com a ideia de contexto. A distribuição dos pesos atribuídos às tarefas revelam a importância, ou falta dela, de uma tarefa em relação a outra. Na grade de avaliação transcrita no capítulo metodológico (SEEDHOUSE et al, 2014), não há menção à distribuição de pontos, salvo pelos valores (0-9) referentes ao desempenho dos candidatos em cada um dos quatro critérios de avaliação (fluência e coerência; pronúncia; recurso lexical; alcance gramatical e precisão). A ausência de

marcadores específicos para cada questão nos leva a crer que a investigação do desempenho de um candidato é feita de maneira holística. Nenhuma das partes pesa mais que a outra, ainda que a dificuldade de cada uma seja variável.

A investigação da distribuição dos pesos e da organização dos critérios de avaliação é uma ferramenta que também podemos utilizar para investigar a concepção de língua que subjaz à prova, seu construto-alvo e sua validade de contexto por outro ângulo. Ao ler os descritores somente, talvez nos sintamos impelidos a classificar a grade de avaliação do IELTS como reflexo de uma concepção de língua estruturalista.

No entanto, se nos voltarmos para a terminologia presente nos descritores individuais de cada uma das quatro seções da referida grade, é possível notar a presença de elementos funcionais, sociocognitivos, bem como estruturais que entregam pistas “ontogênicas” dos possíveis processos envolvidos no desenvolvimento dessa grade de avaliação na seção 5.5 de análise. A presença desses elementos mitiga uma concepção unicamente estruturalista de “cena” e traz à baila a grade de avaliação e sua pontuação como instrumento de investigação sociocognitivo para a validade de contexto do exame oral. É importante que investiguemos a grade porque, por mais que as tarefas apresentem, em si, características sociocognitivas, não há como os instrumentos alinhados com seus propósitos desse ponto de vista.

Para melhor compreender o conceito de língua que subjaz à grade de avaliação, comparamos e contrastamos dois descritores de cada um dos critérios da grade de avaliação com base em nosso referencial teórico. Então, vamos à análise.

5.5 Análise da concepção linguístico-contextual e dos descritores

Como a grade avaliativa apresenta o que é esperado do candidato para seus resultados, apresentamos uma análise em que comentamos dois descritores de cada uma das seções da grade de avaliação que, no nosso entendimento, contemplam os aspectos mais relevantes para uma inferência da concepção de língua subjacente ao exame. Para tanto, apresentamos o quadro 16, que contempla esses componentes.

Quadro 16 - Excerto de descritores para investigação de construto e contexto

Fluência e Coerência	Pronúncia	Recurso Lexical	Alcance gramatical e precisão
<p>8</p> <ul style="list-style-type: none"> fala fluentemente apenas com repetição ou autocorreção ocasional; a hesitação geralmente está relacionada ao conteúdo e apenas raramente à pesquisa de expressões desenvolve tópicos de forma coerente e adequada 	<p>8</p> <ul style="list-style-type: none"> usa uma ampla gama de recursos de pronúncia mantém o uso flexível de recursos da fala conectada, com apenas lapsos ocasionais é fácil de entender durante a fala; O sotaque da L1 tem efeito mínimo na inteligibilidade 	<p>9</p> <ul style="list-style-type: none"> usa vocabulário com total flexibilidade e precisão em todos os tópicos usa linguagem idiomática com naturalidade e precisão 	<p>9</p> <ul style="list-style-type: none"> usa uma gama completa de estruturas de forma natural e adequada produz estruturas consistentemente precisas além das características de "deslizes" da fala do falante nativo
<p>4</p> <ul style="list-style-type: none"> não é capaz de responder sem pausas perceptíveis e por vezes fala devagar, com repetição frequente e autocorreção vincula frases básicas, mas com o uso repetitivo de conectivos simples e algumas falhas na coerência 	<p>7</p> <ul style="list-style-type: none"> mostra todas as características positivas da Banda 6 e algumas, mas não todas, das características positivas da Banda 8 	<p>6</p> <ul style="list-style-type: none"> tem um vocabulário amplo o suficiente para discutir tópicos extensamente e tornar o significado claro, mesmo que com inadequações geralmente parafraseia com sucesso 	<p>4</p> <ul style="list-style-type: none"> produz formas de frases básicas e algumas frases simples corretas, mas estruturas subordinadas são raras erros são frequentes e podem levar a mal-entendidos

Fonte: (SEEDHOUSE et al, 2014)

As passagens destacadas no quadro anterior evidenciam elementos relevantes para a concepção de língua da grade de avaliação e do exame oral. Por exemplo, na seção de fluência e coerência, há menção à “hesitação”. Hesitar é um sinal de processamento cognitivo de reflexão, uma sofisticada atividade mental superior (VYGOTSKY, 1986), que exige, no caso em tela, o desenvolvimento de muitas competências na língua-alvo. Hesitar para uma busca por expressões idiomáticas mais adequadas ao contexto é um sinal de domínio de estilo e práticas sociais, ainda que inconscientemente. Trata-se de uma consideração ao acionamento dessa função mental superior, um dos pontos nevrálgicos da teoria vygotskiana (VYGOTSKY, 1986). A palavra “adequada” também relaciona-se à concepção de competência interacional de Johnson (2003), que concebe a aquisição da língua como interdependente de mediadores materiais, como pessoas e ambientes sociais, e imateriais, como a própria língua utilizada em

um contexto, para que se entenda os padrões interacionais predominantes num determinado contexto interacional.

Se olharmos para o descritor superior na seção de pronúncia no Quadro 16, em que há menção ao “uso flexível de recursos da fala conectada”, vemos que esse critério pode ser interpretado como um traço linguístico que reflete as origens e/ou experiências do candidato, assim como sua condição como interactante, inclusive considerando a(s) sua(s) comunidade(s) de fala e prática. Não podemos afirmar que um avaliador-mediador deve dominar todos os elementos citados como possíveis evidências da vivência experiencial de um candidato, mas podemos, certamente, considerar o texto do descritor “8”, de pronúncia, como alinhado à concepção de Gumperz (1972), que concebe pronúncia como reflexo de prática social.

Na seção referente ao vocabulário, ainda que já tenhamos apontado para o desalinhamento entre as tarefas do exame oral e as práticas interacionais alinhadas a propósitos acadêmicos, não podemos deixar de conceber que a concepção de “domínio lexical” é tida como um dos vários critérios que Van Dijk (2012) menciona ao discorrer sobre contexto. Ao falar a respeito da adequação, o autor cita aspectos tipicamente tidos como estruturais, mas que, quando elucidados sob sua concepção de contexto, tornam-se claramente interacionais. A adequação, que pode ser lexical, sintática, e até variacional, uma vez que a grade de avaliação aponta para o termo “flexível”, indica considerações sobre ao menos um, ou vários esquemas mentais. A inadequação, por consequência, indica o contrário.

Os aspectos gramaticais da língua são, historicamente, um dos tópicos mais debatidos, senão os mais debatidos dentre linguistas. Apesar de haver claro viés estrutural na grade de avaliação quanto ao aspecto gramatical, não há como interpretá-lo como um critério prescritivista - não há língua sem estrutura. O que achamos particularmente marcante na seção de alcance gramatical e precisão da grade avaliativa do exame oral do IELTS foi a concepção de “deslize” integrada à nota máxima. Um “deslize” comum cometido por falantes nativos é, por exemplo, a utilização da locução verbal “ain’t”, cujas funções morfossintáticas substituem inúmeras outras (CHESHIRE, 1981). A perspectiva gramatical, então, parece estar alinhada à sociolinguística variacionista laboviana, por exemplo, o que gera uma grade de avaliação polivalente e fundamentada em diferentes teorias linguísticas. Essa visão polivalente não

trabalha contra a grade, ao contrário, manifesta-se em favor de uma análise que utilize o mais avançado que cada campo da linguística pode oferecer.

A concepção sociocognitiva de língua, por consequência, está presente em vários elementos da grade, como já discutido nos parágrafos anteriores. No entanto, reiteramos que os elementos interacionais que caracterizam exames acadêmicos, ou gerais, não se manifestam explicitamente, e sim por meio de funções e interações voltadas para a operacionalização de um construto de fala voltado para o uso geral da língua. Desenvolveremos mais como isso pode estar refletido nos descritores da grade de avaliação na análise de insumo.

5.6 Análise dos critérios disponíveis

Segundo Weir (2005), um teste deve ofertar o máximo de recursos disponíveis para que os candidatos ao exame possam se preparar de maneira justa e inequívoca. A (não) familiaridade com os critérios, segundo o autor, afeta a mobilização dos recursos cognitivos dos candidatos e, por consequência, o desempenho na tarefa a ser realizada. Taylor (2011) defende que disponibilizar materiais que elucidem os critérios de avaliação, correção, produção esperada, dentre outros, afeta drasticamente o comportamento do candidato em termos de preparação, apesar de haver pouca evidência literária que aponte para um potencial prejuízo caso recursos não sejam disponibilizados.

Quanto aos critérios disponíveis, muito dos materiais para o público, incluindo acadêmicos e didáticos, são disponibilizados no próprio *website* do exame IELTS. Há também um breve tutorial sobre o teste chamado *Road to IELTS* (Estrada para o IELTS) para aqueles que não têm familiaridade com as etapas e critérios do exame.

Para a realização desse trabalho de dissertação, utilizamos, inclusive, pesquisas que foram financiadas pelo BC, IDP e CAE. O *website* disponibiliza abas que direcionam o olhar do usuário de acordo com seus objetivos. Contudo, critérios como a grade de avaliação não são disponibilizados na íntegra para assegurar a confidencialidade de aspectos característicos do exame em todas as suas instâncias e modalidades.

Por mais que haja limitações de acesso, foi por meio dos critérios disponíveis que conseguimos, aos poucos, investigar cada uma das partes do exame oral do IELTS sob as lentes sociocognitivas de nosso referencial. A disponibilização de material facilita a co-construção de interações em diferentes instâncias, seja nas instâncias mediador-candidato, ou pesquisador-instituição.

5.7 Análise da ordem dos itens

Já previamente citada, a ordem dos itens tem como objetivo sequenciar, de maneira lógica, progressiva e didática, isto é, as questões de um exame normalmente são organizadas em dificuldade crescente e podem, ou não, estarem tematicamente interligadas. Galaczi e French (2011) e Field (2011) apontam, respectivamente, para os benefícios em termos de engajamento com as tarefas e mobilização de recursos cognitivos que um bom sequenciamento de questões pode trazer ao candidato.

Em termos de engajamento, caso o exame se mostre complexo, pode desencorajar o candidato a continuar realizando-o, perdendo, assim, a oportunidade de interagir com os colaboradores e/ou instrumentos de avaliação. Em termos de mobilização cognitiva, caso candidatos de baixo nível de proficiência, porém obstinados em ficar até o final do exame, se depararem com questões complexas no início da avaliação, pode ser que usem seu tempo de maneira inadequada. Bachman (1990) postula que, a rígida observação de condições como duração do exame e ordem dos itens são fundamentais para assegurar uma experiência sistematicamente controlada dentro dos parâmetros contextuais que um exame se propõe a simular. O reflexo de tal comportamento afeta positivamente a validade e confiabilidade do teste.

Na seção seguinte, serão analisadas as questões que norteiam as interações do exame oral do IELTS.

5.8 Análise das tarefas

Para investigar o alinhamento entre a ordem dos itens, demanda das tarefas, e descritores da grade de avaliação, utilizamos uma análise com base em múltiplos conceitos: abordamos, nas tarefas, o tipo de interação utilizada e a dificuldade cognitiva. A análise consta em três tópicos da seção 5.9 e, nelas, retomamos alguns dos aspectos já mencionados nas análises anteriores. O tipo de interação e a complexidade do conteúdo foram organizadas conforme as definições de rotinas interacionais de Bygate (1987), já as funções interacionais suscitadas nas tarefas foram organizadas, mais a frente, na seção referente aos aspectos funcionais do exame. A análise da dificuldade cognitiva foi baseada, primordialmente, nos comentários tecidos por Field (2011) e nas definições presentes no QECR (COUNCIL OF EUROPE, 2001).

Junto dessa análise, fazem-se necessárias observações das características predominantes na estruturação das tarefas. Para tal, organizamos quatro critérios para categorizar os temas elencadas em cada uma delas, a saber: concreto, abstrato, familiar/pessoal e não familiar/impessoal. Por fim, para a confecção do texto explicativo de cada uma das partes da prova, investigamos: sua complexidade, sua estruturação, o papel do candidato e do mediador, sua temática e sua demanda cognitiva, não necessariamente nessa ordem, uma vez que a tarefa não é linearmente organizada nessa disposição, e as interações não são uniformemente estruturadas.

5.9 Análise das limitações temporais

As limitações temporais relacionam-se com o tempo de planejamento disponível aos candidatos para a realização da tarefa. Taylor (2011) aponta que a relação entre planejamento, fluência, precisão e complexidade não são, necessariamente, lineares. O que vemos, tal qual Taylor, como relevante à análise das limitações temporais, é a familiaridade que um candidato tem com a prova e seus esquemas mentais previamente estabelecidos.

Isto é, esquemas mentais refletem a relação experiencial de um sujeito com diferentes ambientes interacionais (ex: vivência profissional, acadêmica, etc.) (VAN DIJK, 2012). Falta de domínio dos esquemas mentais ou da quantidade de tempo necessária para responder uma tarefa pode influenciar a habilidade do candidato de mobilizar e operacionalizar funções cognitivas complexas por meio da fala. Contudo, uma vez que há materiais informativos disponíveis e de fácil acesso pelo *website* do exame IELTS, a familiaridade com os temas e tempo do exame pode ser contornada.

Análise sociocognitiva da ordem das tarefas da parte 1

- As tarefas da parte 1 tratam, primordialmente, de um apanhado de perguntas e respostas majoritariamente estruturado e concreto, com prevalência de informações factuais de fácil previsibilidade, elicitando, poucas vezes, funções ligeiramente mais complexas da linguagem (BROWN; YULE, 1983).
- Dados os benefícios citados por Bachman (1990), Galaczi e ffrench (2011) e Field (2011) quanto à ordem das tarefas, a função dessa parte do exame oral nos parece ser voltada para gerar conforto e mobilizar esquemas cognitivos (VAN DIJK, 2012) previamente existentes na mente dos candidatos.
- O papel do candidato, então, limita-se a interagir com as perguntas feitas, havendo pouco espaço para divagação ou quebra do formato hierárquico instaurado por meio da tarefa (WEIR, 2005).
- A mediação da tarefa, apesar de ter caráter interativo direto (face-a-face), pois firma-se entre a relação criada entre mediador e candidato, parece-nos ser controlada, com características similares às de uma tarefa semi-direta (tarefas computadorizadas ou gravadas), pois a interação do examinador limita-se, segundo as transcrições de Seedhouse et al (2004), a eliciar respostas curtas. Galaczi e ffrench (2011) corroboram a nossa concepção da parte 1, uma vez que consideram reciprocidade um quesito fundamental às interações face-a-face.

Análise sociocognitiva da ordem das tarefas da Parte 2

- Trata-se de uma tarefa de fala estendida, também conhecida como turno longo, habilidade indispensável aos ambientes acadêmicos e profissionais. Os tópicos são, ainda, de cunho familiar, mas exigem maior especulação da parte do

candidato em relação à primeira parte do exame, pois adicionam, por exemplo, juízos de valor, que são características de cunho abstrato e de função opinativa.

- Apesar de ainda ser mediada pelo examinador, a tarefa, em comparação às da parte 1, demanda mais recursos cognitivos do candidato do que a primeira. Além da mescla de tópicos relativamente concretos e abstratos, o candidato deve responder à tarefa de forma que sua fala soe como um texto, e não como respostas individuais a perguntas individuais. Dessa forma, nós a consideramos uma tarefa semi-estruturada.

Análise sociocognitiva da ordem das tarefas da parte 3

- Das tarefas presentes no exame, as da parte 3 são as mais cognitivamente desafiadoras. Elas se relacionam à parte 2 em tema, mas distanciam-se quanto ao caráter da informação. O que antes estava sendo tratado de forma relativamente concreta na parte 2, dentro da esfera de conteúdo de cunho familiar do candidato, passa para a esfera social e torna-se consideravelmente mais abstrato.
- O examinador aparenta, segundo as transcrições de Seedhouse et al (2004), ter mais liberdade para suscitar funções complexas da linguagem, tal como pedir por explicações e justificativas. (BYGATE, 1987; BROWN; YULE, 1983).
- Por suscitar funções mais complexas da linguagem com base nas interações e informações produzidas pelo candidato, como em uma conversa real (IELTS, 2021), a tarefa torna-se aberta, pois apesar de uma série de esquemas mentais terem sido mobilizados na parte 2, as mobilizações cognitivas que se instanciam na parte 3 são co-construídas com base na interação, função e complexidade elencadas pelas perguntas do mediador e respostas do candidato (GALACZI; FRENCH, 2011).

Os esquemas mentais prévios necessários às tarefas da prova, no entanto, dependem de familiaridade temática pré-estabelecida na língua-alvo. Esse é outro problema crucial, apontado por Bachman (1990), que é referente à relação existente entre o tempo disponível para uma questão e a proficiência de um candidato. Diferentes níveis de proficiência com a língua requerem diferentes períodos de processamento. Isso quer dizer que, mesmo que um candidato tenha vasta experiência com diferentes assuntos em sua língua primeira, se tais habilidades não forem cognitivamente construídas na língua-alvo do exame pouco será capaz

de ser operacionalizar em favor de amostras que tornem a análise do construto-alvo favorável à nota do candidato. O contexto de cada língua tem suas idiossincrasias para categorizar o mundo, e aprender uma nova língua requer a reorganização de esquemas mentais.

O IELTS, nesse sentido, apresenta situação passível de questionamento. A variabilidade de níveis cujo exame afere é grande. Relembrando, mais uma vez, que quatro níveis distintos de proficiência podem ser certificados, B1, B2, C1 e C2, pela mesma prova e avaliados pelos mesmos descritores para propósitos diferentes. Se analisarmos as diferenças entre cada um dos níveis pelo parâmetro “quantidade de horas médias de ensino guiado” para a conclusão de um nível, vemos a gritante diferença. Tal análise encontra-se no Quadro 17.

Quadro 17 - Relação de horas de ensino assistido e proficiência

Nível de acordo com o QECR	Exames de proficiência oferecidos pela Cambridge English	Número de horas de ensino assistido
C2	<i>C2 Proficiency</i>	1,000—1,200
C1	<i>C1 Advanced</i>	700—800
B2	<i>B2 First</i>	500—600
B1	<i>B1 Preliminary</i>	350—400
A2	<i>A2 Key</i>	180—200
A1	<i>A1 Starters/Movers</i>	90-100

Fonte: (CAMBRIDGE, 2021).

Segundo Wigglesworth (1997) e Galaczi e ffrench (2011), normalmente, porém, de maneira potencialmente equivocada, considera-se que o aumento do tempo de planejamento pode auxiliar candidatos, ou mesmo alunos, cuja proficiência com a língua-alvo é baixa, a formular respostas mais complexas. No entanto, Wigglesworth (1997) demonstra que a afirmativa não se sustenta para usuários de baixa proficiência. Ao contrário, os falantes mais beneficiados pelo aumento do tempo foram os de alta proficiência durante a simulação de um contexto sociocognitivo de natureza complexa. A segunda pesquisa de Wigglesworth (2001)

demonstra que o aumento do tempo pode, na verdade, piorar o desempenho de falantes de baixa proficiência, ainda que em tópicos familiares.

Creemos que um dos motivos relaciona-se à provável limitação (em seu sentido linguístico) cognitiva de lidar com as informações que o usuário reuniu em sua mente, pois lhe faltam recursos para organizar (uma competência) e expressar (outra competência), o que foi experiencialmente angariado. Dessa forma, os apontamentos de Galaczi e French (2011) estão em total consonância com as pesquisas de Wigglesworth (2001), uma vez que as autoras sugerem que o planejamento deve ser disponibilizado somente em tarefas cognitivamente complexas para usuários proficientes, ou em que o contexto assim exige (ex: contexto acadêmico de apresentação de trabalhos).

Planejamento, ou a falta dele, são características naturais à fala (SKEHAN, 1998). A importância e presença de estratégias de planejamento depende dos contextos e dos esquemas mentais prévios que podem vir a permear uma determinada interação oral. Conseqüentemente, um exame de proficiência deve, da maneira mais alinhada possível, trazer isso à tona. No caso do exame IELTS, o tempo disponibilizado na parte 2 licencia a ativação de esquemas cognitivos que justificam a falta de preparação temporal para as perguntas complexas da parte 3. A conexão temática, segundo Weir, O'Sullivan e Horai (2006), é uma maneira de tornar a comunicação que se instância dentro de um exame, assim como o uso do tempo disponível para a fala, contextualmente situada. Na parte 1, em que não há preparação, o postulado dos três autores, assim como os comentários de Skehan (1998), também se mostram verdadeiros, uma vez que os assuntos são de natureza familiar e concreta, ou seja, fazem parte do nosso cotidiano e são de fácil acesso cognitivo. A seguir, apresentamos um quadro resumo com a natureza e demanda cognitiva das tarefas, assim como interação e tempo disponível para cada parte.

A literatura que utilizamos aponta para uma relação entre tempo e desempenho; no entanto, concebemos que as informações, além de vastas, são controversas. Logo, é notável que as condições simuladas em um exame são, normalmente, mais breves e simples do que suas contrapartes reais e isso pode tornar a simulação e necessidade por limitações temporais desnecessárias (GALACZI; FRENCH, 2011).

Quadro 18 - quadro resumo das demandas cognitivas médias por tarefa

	PARTE DA PROVA ORAL DO IELTS	TIPO DE TAREFA/MEDIAÇÃO	NÍVEL CARACTERÍSTICO
MENOR DEMANDA COGNITIVA ↑ gradação de demanda cognitiva ↓ MAIOR DEMANDA COGNITIVA	PARTE 1	ESTRUTURADA E CONTROLADA PELO MEDIADOR/AVALIADOR	A2/B1
	PARTE 2	SEMI-ESTRUTURADA PELO MEDIADOR/AVALIADOR	INTERMEDIÁRIO B1+
	PARTE 3	ABERTA COM PARTICIPAÇÃO MEDIADOR/AVALIADOR	INTERMEDIÁRIO B1+

Fonte: QECR (2001)

No entanto, Weir (2005) menciona que sempre que considerarmos longos turnos de fala como uma possível tarefa de um teste, devemos pensar avidamente nas condições reais de operacionalização de contexto para que elas reflitam a escolha do tempo de preparação necessário. O autor menciona, especificamente, um exame do Instituto de Linguistas, que apesar de não ser de proficiência linguística geral, disponibilizava meses para que os candidatos pudessem preparar uma apresentação satisfatória e, invariavelmente, era a parte do teste em que todos os candidatos obtinham melhor desempenho.

A referência às discussões sobre limitações temporais que devemos impor relacionam-se também, de outro ponto de vista, com as relações entre validade de contexto e de conteúdo. Como argumentado anteriormente, o exame oral do IELTS não se adequa diretamente a algumas das funções e demandas cognitivas normalmente presentes em contextos universitários, o que afeta, parcialmente, sua validade de contexto. A discussão sobre conteúdo, no entanto, permanece pouco estudada, segundo Inoue et al (2018).

De um ponto de vista ensaístico, cremos que pequenas diferenças modulares podem ser feitas de acordo com os as situações em que os candidatos procuram se inserir e, com base nisso, diferentes funções, ou tarefas, consideradas. Como Weir (2005) coloca, recursos tecnológicos trabalham em favor do universo de elaboradores de exames de proficiência, e pequenas alterações funcionais, nas vias da discussão proposta por Luoma (2004), certamente podem direcionar a produção de exames mais contextualmente alinhados.

5.10 Análise de insumo

As questões referentes ao insumo unificam vários dos critérios de análise importantes à validade de contexto que o exame simula (WEIR, 2005). O primeiro critério do modelo, o canal de comunicação, possui quatro categorias.

Quadro 19 - Tipos de insumo

Auditivo	insumo fornecido por um examinador ou vídeo/áudio
Escrito	textos disponíveis, normalmente disponibilizados com léxico inferior ao nível avaliado
Visual	Fotos, desenhos, pinturas, etc.
Gráfico	Gráficos de barra, tabelas, etc.

Fonte: (GALACZI; FRENCH, 2011).

Cada um dos insumos suscita respostas únicas em termos cognitivos. A mérito de ilustração, insumos cujo material é visual tendem a ser mais bem aproveitados por candidatos de proficiência mais alta, uma vez que podem explorar mais as propriedades abstratas (ex: sentimentos evocados, ou opiniões particulares sobre problemas sociais) de uma imagem. Há também de se considerar a adequação de um insumo em relação à validade de contexto de um exame. Em exames cujo objetivo é trazer à tona habilidades relacionadas a negócios, por exemplo, gráficos podem ser particularmente úteis e contextualmente alinhados (GALACZI; FRENCH, 2011).

No caso do exame oral do IELTS, há dois diferentes tipos de insumo: o auditivo e o escrito. Não consideramos que o estímulo visual exista no teste, ainda que ele seja feito face-a-face, pois não se trata de uma imagem estática. Ao contrário, o candidato e o mediador encontram-se em situação interacional o que, segundo McNamara (1997), potencializa a

co-construção de discursos bi ou pluridimensionais, dada a necessidade de ajuste discursivo tanto da parte do mediador, quanto do candidato.

Como já constatado na investigação da parte 1 da prova, não a consideramos completamente interacional por ter baixo índice de reciprocidade. Goffman (1967), por exemplo, aponta para a maneira como cada ambiente suscita diferentes rituais de interação. Ao perguntar para alguém, por exemplo, sobre seu dia, seus estudos ou onde um sujeito mora, espera-se, normalmente, mais do que uma outra pergunta subsequente da parte do enunciador inicial. Em interações tipicamente sociais, há expectativa de troca de informações numa interação em que os interactantes se comunicam em “boa fé” (GOFFMAN, 1967), isto é, com a intenção de manter a comunicação.

Compreendemos que, por se tratar de um exame, deve-se limitar o nível de interação possível e controlá-lo para que amostras do construto-alvo possam ser geradas e, por essa razão, não vemos esse aspecto da parte 1 como um aspecto negativo, mas sim, inerente ao contexto em que a interação é estabelecida: um exame de proficiência. Assim sendo, classificá-lo, inteiramente, como direto, ou como semi-direto, torna-se difícil tarefa, por não estar gravado/computadorizado, mas ao mesmo tempo, por não permitir que rituais de interação apropriados emergjam.

Chun (2006) corrobora nossa análise do formato da parte 1. A autora considera que tarefas similares mediadas por canais de comunicação diferentes podem afetar a percepção cognitiva dos objetivos de uma tarefa. Por exemplo, ao passo que examinadores não devem, até para não afetar a produção do candidato, reagir interacionalmente, expressando surpresa, ou interesse, a um comentário feito pelo candidato, eles podem, por exemplo, gesticular com as mãos para que ele continue a falar caso ainda tenha tempo restante em uma tarefa.

O tipo de insumo, assim como outros aspectos da prova, ajudam-nos a desvelar construtos paralelos que também são suscitados dentro da prova. Em exames cuja interação entre candidatos configura-se como parte integrante, há de se considerar que “interação falada” é um dos construtos operacionalizados. Em exames computadorizados, por exemplo, o construto é avaliado pela complexidade do desempenho operacionalizado pelo candidato (GALACZI; FFRENCH, 2011; FIELD, 2011), então, a qualidade (mensurada pelos descritores) da função cognitiva mobilizada é o construto mais importante.

O IELTS, nessa instância, estabelece diferentes construtos em relação à fala. Além das funções já elencadas em Bygate (1987), Brown e Yule (1983) e QECR (COUNCIL OF EUROPE, 2001), há o construto “interação falada”, na última parte da prova, que parece estar particularmente ligada ao critério “fluência e coerência” da grade de avaliação (SEEDHOUSE et al, 2014). Um dos critérios estabelecidos, por exemplo, é o da nota 6, em que o descritor procura servir de base para que se possa avaliar se o candidato está disposto a falar por longos períodos, embora possa perder coerência às vezes, devido à repetição ocasional, autocorreção ou hesitação (SEEDHOUSE et al, 2014). Tal descritor está alinhado, por exemplo, ao modelo de proficiência de Bachman (1990), à competência interacional (JOHNSON, 2003), e também às competências necessárias a contextos e construtos acadêmicos listados por Inoue et al (2018).

Quanto à parte 2, cuja interação é mediada tanto pelo meio auditivo (a fala do examinador), quanto pelo meio escrito (os tópicos a serem desenvolvidos no elemento motivador apresentado em cartão entregue ao candidato pelo examinador) (SEEDHOUSE et al, 2004; IELTS, 2021), também temos considerações.

Há duas instâncias cognitivas diferentes na formatação da tarefa. A primeira é referente às instruções lidas pelo examinador, em que o tempo limite de processamento antes da produção (1 minuto) é dado; os instrumentos disponíveis para a preparação da resposta (papel e caneta) e interações permitidas (tomar notas e narrar a história/evento [isto é, com base na nossa interpretação da parte 2 e com base em Luoma (2004), sugerindo a prática convencional de reprodução de funções e troca de temas]) são explicitadas; o enunciado contextualiza o tema da tarefa após a apresentação das instruções, como já descrito na primeira seção 5.2 de análise do exame oral, e os comandos são apresentados no cartão de resposta, elencando o que deve ser feito em relação ao tema.

As atitudes delegadas ao examinador contextualizam o ambiente, assunto e padrão de interação que o candidato *deveria* tomar. Enfatizamos “deveria”, pois há investigações que demonstram candidatos, por vezes, eximindo-se de interagir de acordo com os padrões de interação instaurados na questão, perguntando, mudando o assunto e/ou interagindo com o examinador durante a avaliação (SEEDHOUSE; MORALES, 2017). Tal observação retoma a natureza da interação ritualística característica dos ensaios de Goffman (1967), mas cremos

que uma investigação da imprevisibilidade interacional da parte do exame oral do IELTS é assunto para uma investigação à parte.

Em breve nota sobre os comandos escritos, Galaczi e French (2011), assim como Field (2011), advertem que a utilização de comandos escritos pode auxiliar o candidato a manter-se dentro da temática e interação estabelecidas pelo mediador-avaliador, inclusive, potencializando o uso de linguagem mais complexa (ex: conectivos com baixo índice de frequência, estruturas sintáticas incomuns à fala espontânea, etc.). O uso de elementos coesivos, por exemplo, provavelmente surge das instâncias temáticas, isto é, os comandos que subjazem à tarefa principal, criados dentro da tarefa da parte 2. Os assuntos suscitados na tarefa, progressivamente, ligam-se e aumentam em dificuldade, suscitando mais recursos cognitivos para conectar uma pergunta da tarefa à outra.

A parte 3 do exame é regida por meio de insumos auditivos, sem apoio escrito para o candidato, uma vez que objetiva simular uma interação falada autêntica. De um ponto de vista sociocognitivo, há, de fato, potencial autenticidade e alinhamento com os propósitos do exame na interação instanciada na tarefa. Isso se dá porque a “interação falada” é crucial tanto nos ambientes acadêmicos, quanto gerais. A complexidade temática presente na parte 3 é variável, como consta em White (2020), e oscila entre tópicos concretos e abstratos, porém, com notável aumento no nível de impessoalidade/familiaridade presente nas questões da tarefa, conforme apresentado em nossa base documental (SEEDHOUSE et al, 2004) e em Bygate (1987).

Em termos de alinhamento com propósitos gerais de comunicação, em nossa análise, as tarefas da parte 3 são as que apresentam mais potencial de mobilizar múltiplos recursos cognitivos, além de exigir de candidatos mais proficientes especulações de potencial cunho acadêmico. Por mobilizar funções de cunho geral, talvez, o exame seja amplamente utilizado para avaliar candidatos em contexto universitário, porém, cremos que o formato do exame não gera oportunidades para interações mais academicamente contextualizadas em termos de conteúdo e função.

5.11 Análise dos modos de discurso

O modo de discurso não deve ser confundido com o canal de comunicação. Segundo Van Dijk (2012), “[O] discurso e suas propriedades podem sinalizar, indiciar ou expressar propriedades dos contextos de muitas maneiras” (p.40, 2012), assim sendo, devemos distinguir as diferenças entre um discurso semântico e outro pragmaticamente situado. Em outras palavras, o modo de discurso refere-se à imensa variedade de esquemas presentes em um exame para operacionalizar discursos pragmaticamente estabelecidos (SHAW; WEIR, 2007).

Escolher um discurso a ser suscitado em consonância com seu contexto original é uma parte da composição da tarefa que pode garantir ao exame maior validade de contexto. Vários modos discursivos já foram elencados ao longo deste trabalho, principalmente sob a ótica de Bygate (1987), Brown e Yule (1983) e do QECR (COUNCIL OF EUROPE, 2001), mas, com entendimento complementar ao que conceitua Van Dijk (2012), posto que, aqui, o termo discurso ora trata de tipologia textual, ora de componentes textuais dessa tipologia, o que na concepção vandjikiana, que é a da Análise de Discurso Crítica, apresenta-se como uma série de elementos concretizadores de “vozes” históricas, coletivas e ideológicas. Para as análises que encontram-se a seguir, fazemos uma breve retomada das funções informacionais (concretas) e avaliativas (abstratas) de Bygate (1987), e de Brown e Yule (1983), que organizam as funções interacionais do discurso, em ordem de dificuldade, da seguinte forma: narração, descrição, exposição, especulação, comparação e argumentação/persuasão.

Ainda relacionado a isso, podemos elencar uma definição que complementa a ideia de discurso, que é a de Brown, Anderson, Shillcock e Yule (1984). Os autores, no contexto de exames de proficiência, avaliam três situações: estáticas, dinâmicas e abstratas. A primeira é relativa à descrição, a segunda à troca de informações e a terceira à divagação abstrata. Ademais, os autores também sugerem que quanto maior for o número de interactantes, tarefas e interações num exame, mais cognitivamente desafiador ele se torna. Organizamos, no Quadro 20, um resumo que descreve a demanda discursiva das tarefas do exame.

Nota-se, a partir da organização do Quadro 20, que os elementos que gradativamente dificultam as tarefas do exame encontram-se nas funções discursivas suscitadas nas tarefas,

nas instâncias de interação mediador-avaliador-candidato variantes, assim como nas demandas cognitivas de organização e construção de amostras que, como já discutido, afasta-se do domínio concreto para o abstrato ao longo das tarefas do exame. Ainda há de se considerar que as tarefas passam a exigir mais das habilidades organizacionais dos candidatos ao longo do exame quanto à natureza do conteúdo. Por exemplo, na parte 1, os candidatos devem lidar com a reprodução de informações e, muito provavelmente, acessar conhecimentos solidificados. De acordo com Bereiter e Scardamalia (1987) e Vähäpääsi (1982), atividades que lidam com conhecimentos pouco hipotéticos, opinativos, ou especulativos, encontram-se na esfera da reprodução de conhecimento, o que os torna cognitivamente simples.

Em contrapartida, tarefas que exigem transformação de conhecimento (especulação e formulação de hipóteses), assim como organização (narração), são de ordem cognitiva mais complexa. Tais habilidades são tipicamente gerais e presentes em inúmeras áreas humanas, e servem, também, a propósitos e contextos educacionais (ex: educação escolar, universitária, técnica, etc.). São, em nossa visão, uma descrição em minúcias das funções mentais superiores descritas em Vygotsky (1986) adaptadas à nossa investigação sociocognitiva.

Quadro 20 - análise sociocognitiva da demanda discursiva

Partes do exame oral	Características	Elementos intensificadores de complexidade
Parte 1	Estática, dinâmica e abstrata Funções discursivas suscitadas: descrição, exposição, especulação e argumentação	Número de participantes: 2 Número de elementos: 11 perguntas dispostas uma a uma, normalmente suscitando uma função discursiva por vez Número de interações: 1 Unilateral (Ouvir e responder)
Parte 2	Estática, dinâmica e abstrata Funções discursivas suscitadas: descrição, exposição, especulação e argumentação	Número de participantes: 2 Número de elementos: 5 perguntas dispostas simultaneamente suscitando múltiplas funções discursivas de uma vez Número de interações: 1 Unilateral (Ouvir e responder)
Parte 3	Estática, dinâmica e abstrata Funções discursivas suscitadas: descrição, exposição, especulação e argumentação	Número de participantes: 2 Número de elementos: 4 perguntas dispostas uma a uma, suscitando mais de uma função discursiva por vez Número de interações: 1 bilateral (ouvir, responder e responder à reação do mediador-avaliador)

Fonte: (BROWN; YULE, 1983; BROWN; ANDERSON; SHILLOCK; YULE, 1984)

Ainda sobre a ordem das questões, pressupomos que as demandas discursivas são utilizadas como instrumento intra-avaliativo para evidenciar usuários cujas habilidades encontram-se em nível mais alto ou mais baixo. Se analisarmos o Quadro 20 novamente,

com o propósito do insumo e que, durante muito tempo, pensou-se que escrita e fala desenvolviam-se de maneira cognitivamente idêntica. O autor aponta, no entanto, que já há evidências neurológicas e sociocognitivas que apontam que o contexto situacional e cultural de produção (VAN DIJK, 2012; MATTHIESSEN; HALLIDAY, 2009), isto é, o ambiente em que a interação social se situa (conversa em um bar *versus* palestra) e o meio de verbalização da interação (escrita *versus* fala), têm peso significativo no processamento cognitivo da compreensão e produção da mensagem. Isso se dá, primeiramente, porque, diferentemente da escrita, que não só requer como também recebe mais oportunidades de planejamento, a fala *caracteriza-se* por desvios de toda sorte. Segundo Hughes (2002), características comuns à fala espontânea são inúmeras hesitações, abandono de argumentos, erros de concordância, lapsos de memória, repetição de vocabulário, brevidade e sentenças inacabadas.

À luz dessas descrições, o conceito de *rotina interacional* de Bygate (1987) diferencia a fala da escrita justamente por considerar que o grau de imprevisibilidade que acompanha interações orais é infinitamente mais presente. Não se trata de um domínio da língua, mas sim do meio em que ela é utilizada como ferramenta. Weir (2005) refere-se a esse conhecimento como *conhecimento rotineiro*, que envolve habilidades de previsão, planejamento cognitivo diferenciado, interação assistida, dentre outros elementos. Em resumo, apesar de ambas - fala e escrita - serem humanamente convencionadas e, por vezes, servirem a propósitos confluentes, na fala encontram-se mais elementos característicos, de maneira simplista, do *improviso* e de interações breves.

Creemos que, por tal razão, como já citado, o enunciado de tarefas tende a ser mais curto, mesmo em tarefas cujo propósito é acadêmico (ex: faça uma apresentação sobre...). Galaczi e French (2011) ressaltam que em especificações utilizadas para a confecção de exames orais, há um limite máximo de palavras por enunciado, e que esse número varia de acordo com o nível de proficiência sendo avaliado, tendo como limite máximo em níveis básicos, 30 palavras. O enunciado mais longo do IELTS encontra-se na parte 2 em decorrência do preâmbulo lido pelo mediador. As demais seções apresentam menos de 30 palavras, e possuem léxico simplificado, pertencente aos níveis A2/B1 (COUNCIL OF EUROPE, 2001).

5.13 Análise da natureza da informação

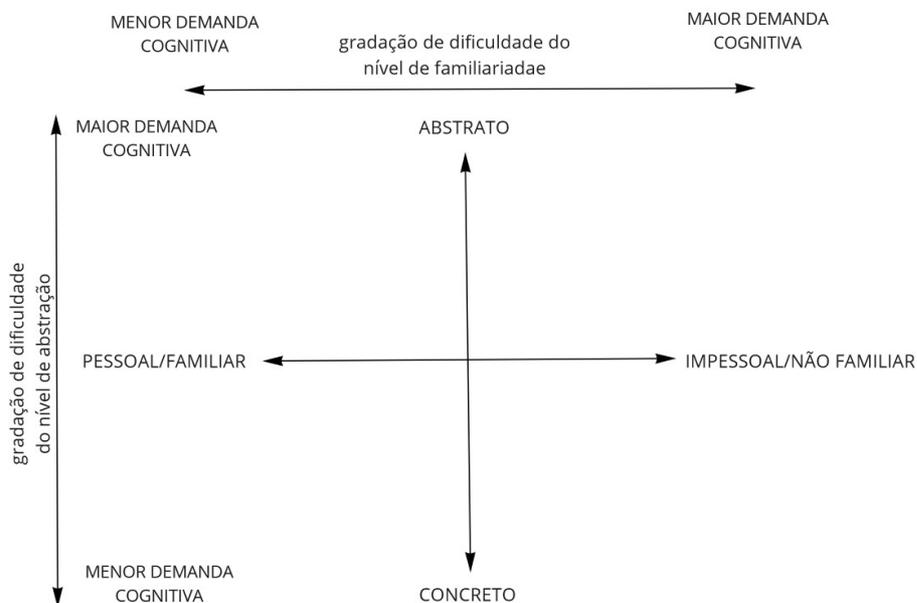
A natureza da informação, como já discutido em algumas seções prévias, refere-se ao quão concreta, ou abstrata, uma tarefa e os construtos que ela operacionaliza são. Podemos pensar na natureza da informação por meio de eixos.

O eixo abstrato-concreto refere-se ao quão observável um evento no dia-a-dia é, enquanto o eixo pessoal-impessoal dita o nível de familiaridade de um assunto em relação aos potenciais referenciais de mundo de um interactante. Em termos gerais, podemos alocar as partes do exame oral do IELTS nas Figuras 5 e 6 da seguinte maneira:

A parte 1 do exame é, majoritariamente, voltada para assuntos referentes a experiências com baixo índice de opinião e mescla de funções e processos cognitivos, o que nos leva a classificá-la como concreta e familiar. A parte 2 do exame possui mais funções discursivas conjugadas, além de ser a tarefa mais exigente em termos de organização cognitiva por mobilizar, além das funções discursivas, concatenação de ideias. Por vezes, quando concatenando ideias, podemos “perder o fio da meada”, ou, como descrito em Hughes (2002), abandonar o argumento que tínhamos formulado.

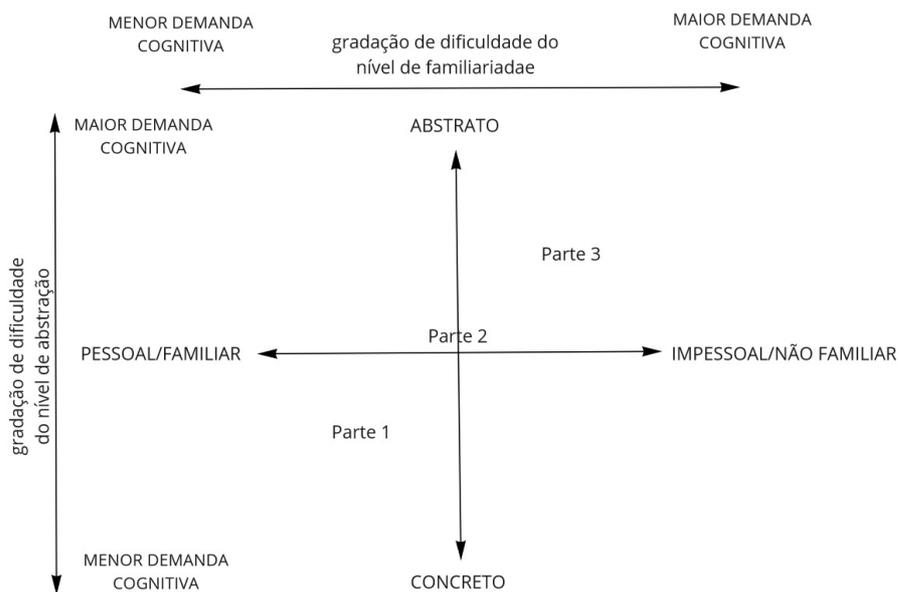
No entanto, tematicamente, por se tratar de uma experiência pessoal, não se afasta do eixo concreto e, simultaneamente, não se aproxima tanto do abstrato. Por fim, a parte 3 apela para funções discursivas voltadas à especulação e argumentação de maneira conjugada sem tempo de processamento; as palavras podem ser familiares aos candidatos, mas os conceitos suscitados são distantes, impossibilitando-os, teoricamente, de desenvolver um argumento que permaneça somente na redoma de sua própria vivência experiencial. Para nós, o traço que mais aproxima a parte 3 do topo dos eixos abstratos e impessoais é o fato de que cenários hipotéticos são postos para que o candidato os avalie, o que exige alteridade e afastamento da realidade.

Figura 5 - gradação de dificuldade temática



Fonte: (GALACZI; FFRENCH, 2011; COUNCIL OF EUROPE, 2001; IELTS, 2021)

Figura 6 - alocação das tarefas em dificuldade temática e cognitiva



Fonte: (GALACZI; FFRENCH, 2011; COUNCIL OF EUROPE, 2001; IELTS, 2021)

A falta de contexto (ex: situações especulativas imprevisíveis) resultante das tarefas mobiliza funções cognitivas superiores e, conseqüentemente, pode prejudicar candidatos cujo nível de proficiência é mais baixo.

5.14 Análise de conhecimento do conteúdo

O conhecimento do conteúdo, principalmente quanto aos temas, segundo Weir (1993) e Hughes (2003), para avaliações cujo alvo é um público heterogêneo, deve ser focado em “conhecimento neutro”. Apesar de acharmos neutralidade um conceito paradoxal, o que os autores em questão querem dizer com neutro é um assunto, ou conteúdo, relativamente familiar a diferentes públicos, faixas etárias e idades. Alderson (2000) também defende o argumento da neutralidade temática ao sugerir que tal prática favorece a mobilização de recursos cognitivos para a formulação de respostas. As afirmativas nos parecem plausíveis, pois menos tempo seria gasto em processos cognitivos de rememoração e mais tempo poderia ser gasto para a produção. Contudo, a visão dos três autores citados aplicar-se-ia, tão-somente, para propósitos gerais e, inevitavelmente, sujeitos com maior domínio de um dos possíveis assuntos da prova pode apresentar melhor desempenho do que outros; por exemplo: pelo acesso ao léxico, ou locuções específicas mais apropriadas para a situação geral elencada durante o exame, um candidato seria capaz de notas mais altas, tal como descrito nos critérios 8 e 9 dos “recursos lexicais” da grade de avaliação. Discutiremos os aspectos referentes à grade de avaliação do léxico mais adiante.

Seguindo as máximas de Weir (2005) para a investigação de validade de contexto, pensando no quão alinhado o conteúdo das tarefas deve ser a um propósito, só podemos afirmar que, em relação ao propósito acadêmico do IELTS, tanto em termos de função discursiva, quanto assunto, há desalinhamento. Afinal, contextos acadêmicos não se desenvolvem somente a partir de contextos de proficiência geral, apesar da transferência de habilidades de um ambiente para o outro ser possível.

Para melhor compreender o que se relaciona à escolha dos assuntos das tarefas em diferentes âmbitos, investigamos os processos gerais envolvidos na validação de exames de proficiência *a priori*, isto é, a fase “piloto” de um exame. Uma vez que não podemos adentrar

essa discussão com profundidade quanto ao IELTS com base em nossa literatura, o que descobrimos foi que, de maneira geral, um grupo focal normalmente é utilizado para gerar a amostra que será utilizada para as deliberações dos elaboradores quanto ao conteúdo das tarefas em relação à validade de contexto. Isto é, uma representação arquetípica de diferentes falantes, em diferentes contextos sociais, com diferentes habilidades interacionais, intelectuais, acadêmicas, profissionais, dentre outras é utilizada para modelar o exame (WEIR, 2005; TAYLOR, 2011).

As amostras representativas procuram testar a aceitabilidade dos assuntos, evitando temas considerados tabus, assim como a reação dos candidatos às tarefas apresentadas, tanto em termos de processamento, quanto em termos de desempenho (FIELD, 2011). Essas amostras auxiliam na geração de perfis aptos para realizar as tarefas do exame.

No caso do IELTS, em seu *website*, há, como já dito, uma seção voltada a pesquisadores, contendo diferentes dados sobre o exame. No que tange o perfil médio do candidato, não há informações sobre poder aquisitivo, grau de escolaridade médio ou preferência por modalidade (*general training* - [treinamento] geral ou *academic* - acadêmica). Tão somente, temos acesso ao desempenho dos candidatos por nacionalidade, sexo, ou nota média em relação aos propósitos. Novamente, é um assunto que, para melhor elucidação, carece de uma pesquisa somente para si.

Se tomarmos como base uma interação tipicamente acadêmica, como um turno longo de fala, podemos ter um bom exemplo do porque o conhecimento sobre o perfil dos candidatos pode vir a ser importante. Turnos de fala longos, como já discutidos, são característicos de instâncias interacionais acadêmicas, mas não tão comuns em instâncias interacionais em contextos gerais. Um turno longo acadêmico, como em uma palestra, tende a durar mais do que 2 minutos, podendo variar de 30 minutos, até, quiçá, 300, além de contar com uso de itens lexicais e terminológicos incomuns ao dia-a-dia.

O que procuraríamos traçar, se possível, com base nos dados demográficos aos quais nos referimos anteriormente, seria a relação do perfil intelectual-médio dos candidatos a fim de investigar, ainda que num subtópico relacionado a esta pesquisa, de forma *avant la lettre*, a necessidade de uma versão da prova voltada para ESP (*English for Specific Purposes* - Inglês para propósitos específicos). Ao discutir a elaboração de exames para propósitos específicos, Galaczi e French (2011), Bachman (1990) e Bachman e Palmer (1996) mencionam como

permitido o uso de tópicos mais restritos e mobilizações cognitivas mais voltadas aos respectivos propósitos sociocognitivos do candidato. Afinal, segundo esses autores, o conteúdo deve ser voltado à faixa etária e experiência do conjunto de amostra de candidatos, assim como também deve ser sensível à cultura e perfil dos mesmos. Segundo nossas fontes, somente assim podemos elaborar uma tarefa com assuntos que contribuam para a validade de contexto de um exame.

5.15 Investigação das demandas linguísticas das tarefas e da grade de avaliação

5.15.1 Aspectos lexicais do insumo e da grade de avaliação

Ao discutir os aspectos lexicais do insumo, concentramo-nos na maneira como a complexidade do código das tarefas é organizada. Há forte interface entre outros aspectos da validade de contexto do teste, como propósito do exame, propósito da tarefa, conhecimento prévio do exame, conhecimento prévio dos possíveis assuntos, nível de proficiência a ser avaliado, entre outras. Para discutir aspectos lexicais, normalmente, parte-se de um corpus de amostra para selecionar o quão alinhada a linguagem utilizada no teste está com o interesse e nível de seus potenciais candidatos. Hindmarsh (1980), Waystage (VAN EK; TRIM, 1998), Threshold (VAN EK; TRIM, 1998), Vantage (VAN EK; TRIM, 2001) e o QECR (COUNCIL OF EUROPE, 2001) são exemplos de seleção de corpus, ou critérios, que visam indicar a performance esperada de um candidato em termos de nível proficiência e uso do léxico e um determinado assunto.

Ball (2002) e Weir (2005) comentam sobre a utilização de computadores para seleção e categorização de corpora específicos. Os autores apontam que bases computadorizadas facilitam a junção de amostras para corpora diversas, tanto para fins avaliativos quanto pedagógicos. Há de se conceber, no entanto, que problemas surgem quando falamos de léxico especificamente voltado a uma área. O'Sullivan (2006), considera que listas de vocabulário específico são mais factíveis do que, por exemplo, listas de cunho terminológico de alta proficiência. Diferentemente de Ball (2002) e Weir (2005), no âmbito das avaliações, O'Sullivan (2006) defende que, em consonância com o modelo de exame oral do IELTS,

mesmo para domínios cognitivos específicos, é possível inferir habilidades específicas do uso da língua por meio de modelos amplos de proficiência, bem como matrizes de referência da mesma sorte, se esses forem aplicados em função da análise do léxico produzido pelo candidato.

Apesar de não estarmos completamente alinhados ao pensamento de O'Sullivan (2006), procuramos entender a fundamentação de seu argumento. De maneira geral, uma lista é gerada com base numa análise de frequência. Há de se considerar que somente a frequência com que a palavra ocorre é levada em conta, ou seja, a presença de uma mesma palavra em forma locucional, ou idiomática, não é diferenciada e, portanto, deve ser feita de forma manual. Considerando essa característica da geração automática de um corpus, tal como Khalifa e Weir (2009), cremos que listas lexicais devem estar alinhadas a propósitos pragmáticos e geradas com base nisso, para então, dentro do contexto pragmático investigado, serem classificadas em termos de frequência.

Contudo, tratando-se, em nosso caso, de uma investigação referente à sociocognição e validade de contexto em língua estrangeira, ou segunda língua, torna-se ainda mais difícil de afirmar como o contexto pragmático original do candidato pode interferir na utilização do léxico. Schmitt (2000) sugere investigações léxico-volumétricas, isto é, a relação entre volume de léxico e proficiência, para a realização de tarefas mais adequadas para contextos de L2 ou LE; o autor aponta que o vocabulário médio para realizar necessidades básicas (COUNCIL OF EUROPE, 2001) varia em torno de 2.000 a 3.000 palavras. Novamente, num contexto de L2 ou LE, um repertório lexical contendo funções cognitivas complexas encontra-se em torno de 10.000 palavras (HAZENBERG; HULSTJIN, 1996).

No caso do exame oral do IELTS, o que vemos é um insumo lexical que atende aos quesitos de neutralidade elencados em Weir (1993). Galaczi e French (2011) argumentam em prol da neutralidade lexical e simplificação do léxico envolvido na formulação de tarefas, principalmente quando em exames de uso geral da língua porque, do contrário, ainda que estivéssemos tratando de níveis avançados de proficiência, estaríamos testando mais do que a habilidade de produção oral dos candidatos no contexto de exame; estaríamos adicionando ao insumo auditivo do exame oral um elemento externo ao construto e à complexidade relativa ao construto-alvo nas tarefas, que é a de compreender o léxico das tarefas.

Segundo as autoras, a formulação lexical do insumo deve ser simples, e operacionalizar, por meio da tarefa, o construto-alvo inerente ao propósito do exame. Por se tratar de um exame que avalia mais de um nível de proficiência, conforme a investigação feita na seção de limitações temporais, notamos, por exemplo, que a adaptação a vários níveis vem por meio da mescla de funções e propósitos das tarefas, mas a avaliação do uso lexical das tarefas e da grade de avaliação é feita de outra maneira.

Para começar nossa investigação do léxico, devemos entender o que um exame espera em termos de elaboração de tarefas e também da produção da parte do candidato, porém, numa retomada mais aprofundada do assunto. A visão da análise volta-se especificamente para o material da prova, isto é, as tarefas, a grade avaliativa e listas lexicais. As listas que mencionamos são fontes utilizadas como guias para exemplificação de seleção de léxico e funções de uma tarefa de acordo com um determinado nível de proficiência. Vale ressaltar que listas lexicais voltadas para o C1 e C2 são delegadas, normalmente, ao elaborador das tarefas de um determinado exame (GALACZI; FFRENCH, 2011). Dessa forma, como não temos contato, ou acesso a documentos que demonstrem as escolhas e processos cognitivos que permearam a escolha das funções ou vocabulário suscitado nas tarefas voltadas para os níveis C1 e C2 investigados na avaliação do IELTS, recorreremos a documentos abertos, como o QECR (COUNCIL OF EUROPE, 2001) e aos descritores da grade de avaliação (SEEDHOUSE et al, 2014).

Após traçar as relações entre léxico, função e elaboração das tarefas, apresentamos uma investigação referente ao “se” e “como” a grade de avaliação alinha-se à validade de contexto do exame em nossa concepção sociocognitiva.

Quadro 22 - Expectativa de produção em relação ao nível de proficiência

Referencial lexical (ilustrativo)	Expectativa de utilização do léxico em termos de familiaridade e limitações lexicais
Waystage (VAN EK; TRIM, 1998); QECR (COUNCIL OF EUROPE, 2001;) A2	<p>Característica semântica do léxico informações relacionadas ao dia-a-dia</p> <p>Limitações para elaboração da tarefa palavras simples</p> <p>Tema para elaboração das tarefas e assuntos do insumo e expectativa de produção do candidato</p> <p>Funções sociais</p> <ul style="list-style-type: none"> - conhecer pessoas pela primeira vez

	<ul style="list-style-type: none"> - fornecer informações factuais de tipo pessoal e não pessoal <ul style="list-style-type: none"> - falar sobre a vida cotidiana - falar sobre interesses - falar sobre gostos - falar sobre atividades de lazer <ul style="list-style-type: none"> - falar sobre vida social Funções linguísticas - expor, descrever
<p>THRESHOLD (VAN EK; TRIM, 1998); QECR (COUNCIL OF EUROPE, 2001;) - B1</p>	<p>Característica semântica do léxico</p> <ul style="list-style-type: none"> - além das informações listadas acima, espera-se também informações de assuntos do dia-a-dia <p>Limitações para elaboração da tarefa</p> <p>palavras compostas contanto que cada uma das palavras apareçam adequados e significados transparente e literal (ex: guarda chuva)</p> <p>locuções verbais com significado literal também são permitidas contanto que apareçam separadamente;</p> <p>Tema para elaboração das tarefas e assuntos do insumo e expectativa de produção do candidato</p> <p>funções sociais</p> <ul style="list-style-type: none"> - detalhes sobre a vida pessoal dos candidatos <ul style="list-style-type: none"> - rotinas diárias - gostos - discussões de alternativas - sugerir e responder a sugestões <ul style="list-style-type: none"> - recomendar - negociar <p>Funções linguísticas</p> <ul style="list-style-type: none"> - expor, descrever, argumentar
<p>VANTAGE (VAN EK; TRIM, 2001); QECR (COUNCIL OF EUROPE, 2001;) B2</p>	<p>Característica semântica do léxico</p> <p>além das informações listadas acima, espera-se também interações voltadas para a interação social e uso da língua geral</p> <p>Tema para elaboração das tarefas e assuntos do insumo e expectativa de produção do candidato</p> <p>Funções sociais</p> <p>vocabulário apropriado em relação ao contexto dar e trocar opiniões sobre uma variedade de tópicos familiares</p> <p>Funções linguísticas</p> <p>comparar, descrever, expressar opiniões, concordar/discordar, sugerir, especular avaliar, etc</p>
<p>QECR (COUNCIL OF EUROPE, 2001)⁵¹</p>	<p>Característica semântica do léxico</p> <p>além das informações listadas acima, espera-se que o candidato possa aplicar todo o léxico anterior de forma geral</p> <p>Tema para elaboração das tarefas e assuntos do</p>

⁵¹ Segundo Galaczi e French (2011), a escolha do léxico do C1 e C2 depende de um conselho. Cada exame possui um conselho.

	<p>insumo e expectativa de produção do candidato</p> <p>Funções sociais evidência de familiarização com o vocabulário do contexto para dar e trocar opiniões sobre uma variedade de tópicos familiares, desconhecidos e abstratos</p> <p>Funções linguísticas comparar, descrever, expressar opiniões, concordar/discordar, sugerir, especular avaliar, etc</p>
QEER (COUNCIL OF EUROPE, 2001);	<p>Característica semântica do léxico além das informações listadas acima, espera-se que o candidato possa aplicar todo o léxico anterior de forma geral flexivelmente</p> <p>Tema para elaboração das tarefas e assuntos do insumo e expectativa de produção do candidato</p> <p>Funções sociais evidência de familiarização com o vocabulário do contexto para a flexibilidade dar e trocar opiniões sobre uma variedade de tópicos familiares, desconhecidos e abstratos</p> <p>Funções linguísticas comparar, descrever, expressar opiniões, concordar/discordar, sugerir, especular avaliar, etc</p>

Fonte: (VAN EK; TRIM, 1998a; 1998b; 2000; COUNCIL OF EUROPE, 2001)

De acordo com a organização de expectativa de produção do léxico exposta no Quadro 22, que contém especificidades inerentes a cada nível de proficiência, é possível notar que características referentes aos níveis A2, B1 e B2 são mescladas aos níveis C1 e C2 por meio de um aumento de complexidade temática no exame oral do IELTS, aumento tal encontrado nas funções discursivas suscitadas nas tarefas, parcialmente analisadas anteriormente.

Laufer (1997) e Schimitt (2000) apontam que o número de palavras caracteristicamente relacionadas a um contexto não é necessariamente um indicador de validade e propósito. Essa afirmativa conflui com o formato do exame oral do IELTS. No entanto, para esses autores, o que realmente indica a relação entre dificuldade e o léxico utilizado na prova é a relação que pode ser traçada entre as tarefas e as situações contextuais, co-contextuais (rituais interacionais pertencentes a um determinado contexto [GOFFMAN, 1967; VAN DIJK, 2012]), assim como as relações co-textuais, que são as em que se unem dois ou mais textos, internos ou externos à interação do exame, em função da validade de contexto que o exame se propõe a assegurar. Tal relação pode ser vista, por exemplo, por meio do alinhamento estabelecido entre a grade de avaliação e suas tarefas.

Uma análise de corpus seria necessária para uma investigação da relação existente entre a frequência de itens lexicais diversos e a variação de desempenho de candidatos voluntários. Assim, a partir dos diferentes resultados de um mesmo candidato a diferentes insumos, estabelecer-se-ia um intervalo de confiabilidade que configuraria o nível de variação. A frequência resultante de tal investigação nos daria aporte para tecer comentários sobre a dificuldade dos itens lexicais utilizados nas tarefas do exame, dado que, para alguns pesquisadores, a frequência de utilização de um item lexical pode indicar sua dificuldade (SCHIMITT, 2000; SCHIMITT; CLAPHAM, 2001). Segundo Hawkins e Butterfly (2009), palavras mais simples tendem a possuir maior frequência, o que facilita com que aprendizes de segunda língua possam otimizar o uso das propriedades sociocognitivas de itens linguísticos mais comuns.

Numa análise superficial⁵² do corpus da amostra, notamos que variações verbais, como conjugações referentes a tempos presentes em níveis A2 e B1 (presente simples, perfeito e contínuo, assim como passado simples - considerando léxico como um aspecto da prova que faz interface com o aspecto estrutural) configuraram a maioria dos enunciados das tarefas. Segundo Schmitt (2009), essa prática assegura com que candidatos de todos os níveis possam entender e realizar de maneira eficaz o que foi perguntado nas tarefas, ainda que não tenham vocabulário para respondê-las efetivamente.

Há formações sintáticas relativamente avançadas no enunciado das perguntas, como *embedded questions* - questões embutidas⁵³, mas que não consideramos onerosas ao processamento cognitivo dos candidatos. Essa observação nos leva a crer que a adequação de linguagem do exame é coerente com a variabilidade de níveis que o exame se propõe a avaliar, mas não necessariamente com a modalidade acadêmica, por mobilizar, somente de maneira breve, os recursos e práticas sociocognitivas pertencentes a esse nicho interacional. Tal afirmativa já havia sido parcialmente exposta no início deste capítulo, na primeira análise das funções discursivas presentes nas tarefas, a qual é complementada lexicalmente agora.

⁵² A análise do corpus foi feita em relação ao léxico utilizado nas tarefas e sua repetição. Não fizemos uma análise de corpus estatisticamente rigorosa para entender a frequência com que itens lexicais variavam, ou repetiam-se, dada a escassez de amostras de exames oficiais, completos. Entendemos que isso tornaria a análise mais confiável e, por isso, chamamos de superficial.

⁵³ “Can you tell me where you live?” - você pode me dizer onde mora? - A pergunta conjuga duas diferentes perguntas - “can you tell me” - você pode me dizer? e “where do you live?” - onde você mora?. Ambas são perguntas características dos níveis A1 e A2, e consideramos que sua conjugação, possivelmente, não é um fator diferencial para o processamento da pergunta.

Em relação à grade de avaliação, investigamos se e como o reflexo da análise de superfície do léxico das tarefas manifesta-se em relação ao construto-alvo e contexto simulado. A seguir, transcrevemos, do capítulo metodológico, somente a seção relativa ao léxico da grade avaliação (SEEDHOUSE et al, 2014), a fim de discutir, de acordo com os pressupostos de Van Dijk (2012) e Weir (2005), o alinhamento da grade aos propósitos do exame.

A discussão será conduzida com base nos níveis (2 a 9) e critérios presentes nos descritores, juntamente de grifos, feitos por nós, para auxiliar a divisão e categorização da discussão. Sob a perspectiva sociocognitiva, tentaremos entender como os descritores diferenciam um nível do outro.

Os primeiros níveis a serem analisados são o 9, o 8 e o 7. Começaremos por esses três níveis porque, como brevemente citado, materiais que definem o léxico referente aos níveis C1 e C2 não são ponto pacífico entre pesquisadores da área. Assim sendo, nos propomos a tentar tecer contribuições para a compreensão do que se configura como inerente à validade de contexto nesses níveis de proficiência. Esses três níveis, 9, 8 e 7, configuram-se como equivalentes, respectivamente, aos níveis C2 e C1. O nível 9 seria, comparativamente, um C2 “alto”, enquanto o 8 seria um C2 “baixo”, ou C1 “alto”, e o 7, um C1 “baixo”, ou B2 “alto”. A categorização exata de cada um dos níveis apresentados nos descritores é complexa, uma vez que não há, no exame oral do IELTS, uma pontuação minuciosamente categórica para indicar em qual nível de desenvolvimento linguístico-cognitivo o candidato se encontra em relação ao nível a ser avaliado (Ex: C1 alto = 1100/1200).

Quadro 23 - Análise da seção de “Recurso Lexical” da grade de avaliação

<p>9</p> <ul style="list-style-type: none"> • usa vocabulário com total flexibilidade e precisão em todos os assuntos • usa linguagem idiomática com naturalidade e precisão 	<p>5</p> <ul style="list-style-type: none"> • consegue falar sobre tópicos familiares e desconhecidos, mas usa vocabulário com flexibilidade limitada • tentativas de usar a paráfrase, mas com sucesso relativo
<p>8</p> <ul style="list-style-type: none"> • usa um amplo recurso de vocabulário de forma rápida e flexível para transmitir significado preciso • usa vocabulário menos comum e idiomático habilmente, com imprecisões ocasionais 	<p>4</p> <ul style="list-style-type: none"> • é capaz de falar sobre tópicos familiares, mas só consegue transmitir o significado básico de tópicos desconhecidos e comete erros frequentes na escolha das palavras • raramente tenta parafrasear

<ul style="list-style-type: none"> • usa paráfrase de forma eficaz, conforme necessário 	
<p style="text-align: center;">7</p> <ul style="list-style-type: none"> • usa recursos de vocabulário de forma flexível para discutir uma variedade de assuntos • usa algum vocabulário menos comum e idiomático e mostra alguma consciência de estilo e expressões idiomáticas, com algumas escolhas inadequadas • usa a paráfrase de forma eficaz 	<p style="text-align: center;">3</p> <ul style="list-style-type: none"> • usa vocabulário simples para transmitir informações pessoais • tem vocabulário insuficiente para tópicos menos familiares
<p style="text-align: center;">6</p> <ul style="list-style-type: none"> • tem vocabulário amplo o suficiente para discutir tópicos extensamente e tornar o significado claro, mesmo com inadequações • geralmente parafraseia com sucesso 	<p style="text-align: center;">2</p> <ul style="list-style-type: none"> • produz apenas palavras isoladas ou enunciados memorizados

Fonte: (SEEDHOUSE; EGGBERT, 2004)

O critério comum aos três níveis, 9, 8 e 7, em relação ao que chamaremos de “uso proficiente do léxico” nessa grade de avaliação, nos parece ser a habilidade de trazer à tona vasta variabilidade de assuntos, tanto em termos lexicais, quanto semânticos. Já a variação de proficiência é expressa em instâncias isoladas, são elas: o uso de paráfrases, estilo, adequação, uso de expressões idiomáticas, flexibilidade e precisão.

A primeira coisa que notamos quanto ao critério “uso de paráfrases” é que, se considerarmos que sinônimos são comumente utilizados para produzi-las, trata-se de uma característica mais sintática do que lexical. Podemos, é claro, parafrasear uma informação apenas por meio da alteração das palavras originais por palavras sinônimas, ou lexicalmente próximas (ex: eu amo sorvete; eu gosto de *gelato*). Sobre as características sintáticas, a paráfrase também pode ser usada para esclarecer, ou explicar, uma informação prévia, o que se configura, ainda nesse caso, como uma espécie de *oração subordinada*, ou mesmo, no caso de oração tematicamente relacionadas, como um recurso retórico para sustentar afirmativas já formuladas.

Há também, ainda no âmbito sintático-estilístico, a utilização de elipses, perífrases, anáforas e catáforas para a construção de diferentes tipos de paráfrases por meio de uma retomada “dêitica” do assunto, um dos aspectos discursivos pertencentes ao modelo de Van Dijk (2012). O domínio de cada uma dessas estratégias implica em diferentes habilidades

cognitivas, assim como esquemas linguísticos prévios contextualmente definidos (VAN DIJK, 2012).

Se tomarmos o simples ato de parafrasear como uma construção sintática que designa elementos previamente mencionados por outros atributos (CEGALLA, 2002), devemos explicar o quão complexa ela deve ser, e até mesmo a maneira como a retomada dos atributos pode ocorrer. Concordamos que todos os exemplos elencados anteriormente configuram uma série de variedades complexas de uso da língua e, quiçá, uma variedade de uso acadêmica da língua, mas cremos que a alocação de tal recurso pertence mais à seção de estrutura gramatical, pelo menos em termos da investigação da frequência com que um determinado candidato pode apresentar variação em sua fala, do que pelo simples ato de trazer uma paráfrase à tona.

Nos é evidente que o número e variabilidade de amostras sintáticas que o candidato for capaz de produzir pode servir como indicador de sua habilidade de parafrasear flexivelmente (9), eficaz e necessariamente (8), ou somente, eficazmente (7). O descritor 9, em nossa percepção, resume um usuário capaz de utilizar todas as diferentes formas que uma paráfrase pode ter. Dito isso, resta-nos investigar como o sentido suscitado pela palavra “necessariamente”, no descritor 8, assim como o da palavra “eficazmente”, no descritor 7, se relacionam e diferenciam.

A palavra “necessariamente”, em termos sociocognitivos, parece-nos estar relacionada à descrição de contexto como construto subjetivo dos participantes (VAN DIJK, 2012). O autor observa que situações interacionais influenciam o discurso por meio das interpretações que atribuímos a elas. A necessidade, ou não, de complemento, depende das interpretações feitas pelos interactantes. Conseqüentemente, a necessidade de paráfrase depende tanto do examinador, quanto do candidato examinado, sem esquecer, é claro, do contexto simulado. Assim sendo, o que torna a paráfrase necessária? Como o contexto relaciona-se à necessidade de uma paráfrase? Ademais, a necessidade ou não de uma paráfrase não depende, também, das experiências prévias do mediador-avaliador?

Se analisarmos a necessidade de paráfrase do ponto de vista das tarefas, é possível notar que somente as tarefas mais abstratas podem evocar a necessidade de explicações, ou desenvolvimento de um determinado assunto de maneira aprofundada, para que emergjam paráfrases que se justifiquem com base no tema e contexto abordados. Dessa forma, a

paráfrase necessária, não seria também eficaz? E sendo ela eficaz, não seria ela, obrigatoriamente, necessária ao contexto? Vemos certa incongruência nesses dois termos utilizados pela grade de avaliação, uma vez que os contextos subjetivos subjacentes à tarefa e pertencentes aos contextos cujo exame procura simular nos levam a ir além da referida grade para que sejamos capazes de justificar, ou não, as diferenças entre um uso “necessário”, e um uso “eficaz” de paráfrases.

Os critérios que diferenciam “estilo”, “adequação” e “uso de expressões idiomáticas”, em nossa visão, encontram-se interrelacionados, amarrados, tal como Weir (2005) preconiza, por um fio temático definido pelo contexto. Essa concepção de estilo manifesta-se no exame por meio de tarefas e descritores que evidenciam sua validade de contexto. Notavelmente, a palavra “estilo” aparece somente no descritor 7, o que nos causa inquietação.

Estilo, segundo Cappella (1996), é o alinhamento estilístico à situação social que os interactantes estão envolvidos, e tal alinhamento pode ser observado por meio da fala, registro, vestimenta, comportamento, origem, dentre outros fatores. Aqui, inferimos o que interpretamos ser o “estilo lexical acadêmico”. Com base no exposto sobre a natureza do léxico e a análise de frequência descrita em Schimitt (2000), Schimitt e Clapham (2001), e Hawkins e Butterfly (2009), que defendem que palavras mais simples tendem a apresentar maior frequência, inferimos que o léxico acadêmico, particularmente o de pós-graduação, encontra-se fora do escopo do léxico considerado geral e cotidiano. Nossa afirmativa é respaldada não só pela análise das funções descritas em Ducasse e Brown (2011), mas também pelas observações de variação de registro feitas em Inoue et al (2018), pois notáveis itens lexicais tendem a aparecer com mais frequência em seminários e palestras do que em ambientes interacionais contemplados dentro do escopo geral de proficiência.

Assim sendo, a definição da concepção de “estilo” listada na grade de avaliação torna-se tão passível de questionamento quanto os conceitos de “necessidade” e “eficácia”, apresentados nos parágrafos anteriores. Não seria o conceito de “estilo”, tanto em termos de função, quanto em termos de contexto, variável? Em nossa concepção, sim, uma vez que interactantes que transitam entre contextos gerais e acadêmicos, apesar de, possivelmente, compartilharem um mesmo idioma como ferramenta de interação, certamente não utilizam os mesmos registros, expressões idiomáticas e, conseqüentemente, comportamentos, em cada uma dessas situações.

O conceito de “adequação” refere-se, em nossa interpretação, ao ritual de interação (GOFFMAN, 1967) que enquadra (GOFFMAN, 1986) um determinado contexto social. Tal afirmativa também é verdadeira para as expressões idiomáticas. Por exemplo, a expressão idiomática *to throw somebody in at the deep end* - algo como, literalmente, lançar alguém em direção às profundezas - que é uma expressão idiomática da língua inglesa tipicamente utilizada em contextos profissionais, para descrever situações em que sujeitos são dados novas responsabilidades sem preparo, ou auxílio, para as tarefas que terão que designar (CAMBRIDGE, 2021).

Avaliar a adequação do uso do léxico de uma expressão idiomática como essa requeriria, principalmente da parte do mediador-avaliador, domínio de esquemas mentais pertencentes ao contexto profissional de alta proficiência (B2, C1 e C2) do uso da língua inglesa. O ato de avaliar “adequação” junto do “uso de expressões idiomáticas”, assim como “estilo”, encontra-se ligado a categorias sociocognitivamente compartilhadas que dependem de uma base cultural e experiencial (VAN DIJK, 2012). Do contrário, julgar “adequação” tornar-se-ia uma tarefa especulativa, caso houvesse falta de conhecimento sobre a expressão utilizada, ou experiencial, caso houvesse conhecimento da expressão utilizada, de alta exigência sociocognitiva, uma vez que o mediador-avaliador seria o responsável por categorizar o desempenho do candidato como adequado ou inadequado. Tal problema dependeria de uma lista lexical e extenso treinamento para que os avaliadores se tornassem capazes de prover pontuação alinhada ao desempenho de acordo com o contexto simulado.

Nesse sentido, entendemos, com base em Schimitt (2000), Schimitt e Clapham (2001), e Hawkins e Butterfly (2009), que uma produção contendo léxico de menor frequência pode levar o mediador-avaliador com vivência experiencial a precisas alocações contextuais; portanto, a utilização do mesmo modelo de prova para dois propósitos diferentes, o acadêmico e o geral, não parece, em sua grade de avaliação, levar em conta a produção do que poderia ser considerado léxico incomum, ou sofisticado, para cada um dos propósitos que se propõe a avaliar. No entanto, reiteramos que para uma afirmação categórica sobre nossa suspeita necessitaria de uma investigação dos documentos e procedimentos de treinamento dos avaliadores para melhor sustentar-se.

Resta-nos investigar os critérios de “flexibilidade” e “precisão”. Como já citado em nossa investigação do que poderia ser considerado paráfrase, entendemos “flexibilidade”

como competência para mobilizar diferentes recursos cognitivos a fim de expressar, ou sustentar, opiniões, por meio de diferentes manifestações lexicais gramaticalmente adequadas e variadas. Consideramos, agora, as possíveis diferenças de qualidade de uso e avaliação deste critério com base nas palavras que as qualificam na grade de avaliação.

“Flexibilidade”, nos descritores 9, 8 e 7, recebe três qualificações: total, ampla ou, somente, “usa recursos de vocabulário de forma flexível”. Considerando o significado de “total”, o que nos soa, destarte, hiperbólico, retomamos alguns dos comentários utilizados para a investigação das possíveis habilidades de um candidato de produzir paráfrases. Isto é, recursos lexicais sendo utilizados com “flexibilidade total”, pelo menos de um ponto de vista sociocognitivo, implicariam numa situação em que seria possível avaliar o alinhamento entre as palavras utilizadas, o contexto simulado e a capacidade cognitiva do avaliador de julgar a amostra com base em seu conhecimento prévio. Só assim poderíamos avaliar o quão válido o contexto evocado pela prova é em relação ao que ela se propõe a avaliar e, por consequência, a variabilidade pragmática potencialmente contida no critério “flexibilidade”.

Dessa forma, nos parece que, no exame oral do IELTS, a avaliação do desempenho do uso de recursos lexicais de um candidato, apesar de aliar a concatenação de ideias de forma coerente junto de léxico variável, não parece estar alinhada à situação proposta. Afinal, a avaliação é humana, ou seja, tal como um computador, mas também diferentemente de um, possui banco de dados (recursos cognitivos) limitado (porém, não necessariamente padronizado em termos de listas lexicais); ademais, as tarefas são voltadas para atos comunicativos relacionados à competência oral generalizada da língua inglesa, o que nos permite inferir que o desempenho produzido não tende, pelo menos com base nas temáticas e funções possivelmente suscitadas (LUOMA, 2004; WHITE, 2020), a garantir validade de contexto.

Nossa conclusão parcial é que, o candidato cujo propósito é acadêmico será avaliado por meio de uma mobilização de recursos inerentemente voltados a uma teoria de proficiência cujo construto é de cunho geral e, conseqüentemente, suas habilidades acadêmicas serão estabelecidas com base na suposição de que há capacidade de transferência e adaptação do construto pertencente a um ambiente geral para um ambiente acadêmico. Seria presunçoso de nossa parte afirmar que não é, e como não temos interesse em especular sobre a capacidade

cognitiva de transferência e alocação de recursos mentais de um candidato de um ambiente a outro, limitamos nossos comentários à validade de contexto do exame.

Quanto aos itens 6, 5, 4, que também são certificados pelo exame, vemos alinhamento entre os critérios organizados na grade de avaliação e a bibliografia citada para definir o uso do léxico nesses níveis, que são, respectivamente, B2 “baixo” ou B1 “alto”, B1 “alto” e B1 “baixo” ou A2 “alto”. Logo, continuamos a defender que, mesmo em casos em que editais acadêmicos exijam desempenho equivalente à nota 6, ou mesmo inferiores, a expectativa quanto ao uso de recursos lexicais deve ser definida e organizada com base no que um candidato acadêmico dos respectivos níveis devesse ser capaz de fazer por meio da língua em seu respectivo contexto.

Com base na análise dos critérios anteriores e no fato de que o exame se propõe a avaliar e alocar múltiplos níveis de proficiência em um descritor, concluímos a presença de ambiguidade nos termos da grade de avaliação e potencial desalinhamento a diferentes contextos sociais. Na próxima seção, faremos a análise dos critérios estruturais da prova.

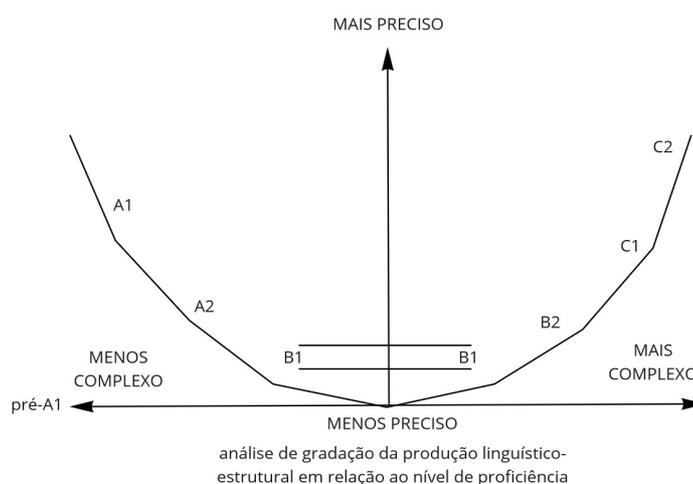
5.15.2 Análise de aspectos estruturais do insumo e da grade de avaliação

Os aspectos estruturais do insumo de uma prova afetam a maneira como as demandas linguísticas da tarefa são percebidas. Ao lidar com os aspectos estruturais do insumo, um elaborador de tarefas deve ter como objetivo traçar relação entre as estruturas características que usuários de cada nível de proficiência conseguem produzir e categorizá-las, para que tarefas apropriadas sejam feitas para cada nível. Galaczi e French (2011) citam que o processo mais comumente utilizado por elaboradores para a confecção de tarefas em que o insumo é apropriado para cada nível é, tal como para a escolha do léxico apropriado, baseado na escolha de um referencial que elenque o que é, ou deixa de ser, apropriado para cada nível.

De acordo com os dados contidos no QECR (COUNCIL OF EUROPE, 2001) e Galaczi e French (2011), candidatos A1 e A2 possuem controle limitado da estrutura da língua-alvo, sendo característico do A2/B1 mais erros de natureza sintática de maneira

sistemática (ex: um erro quanto à conjugação da terceira pessoa do singular). Em candidatos de nível B2, o nível de controle gramatical é relativamente maior, com maior repertório de formas verbais (ex: presente e passado simples, aspectos contínuos e perfeitos) e estruturas fundamentais (ex: ordem das palavras, formulação de perguntas, uso de verbos auxiliares). Nos usuários de língua cujo nível é C1 ou C2, o esperado é que eles sejam capazes de manter o controle gramatical de suas estruturas com precisão praticamente constante junto de recursos sintáticos que os permitam produzir enunciados de maneira precisa e apropriada, incluindo sensibilidade ao registro necessário à interação. Para melhor aporte visual da explicação, adicionamos a Figura 7, que exemplifica a relação ilustrada.

Figura 7 - relação entre proficiência, precisão e conteúdo⁵⁴



Fonte: (PICCARDO; NORTH, 2019)

Tal como nas seções anteriores, focamos nossa análise no que se espera em termos de produção, primariamente, para depois investigar a maneira como a grade de avaliação do exame oral do IELTS contempla a produção de seus candidatos. Consideramos, como aspectos estruturais do exame, as seções de “pronúncia” e “alcance gramatical e precisão” da grade. Para fazer a descrição do que se espera em termos de produção gramatical de candidatos dos níveis A1 ao C2, utilizamos, primariamente, Taylor (2011), Waystage (VAN EK; TRIM, 1998a), Threshold (VAN EK; TRIM, 1998b), Vantage (VAN EK; TRIM, 2001) e

⁵⁴ As linhas que ligam os dois marcados como nível B1 são indicativos da liquidez característica das habilidades do nível.

o QECR (COUNCIL OF EUROPE, 2001). Associamos as definições de cada um desses referenciais para poder definir o que é esperado de cada um dos níveis de proficiência que são contemplados pelo exame e grade de avaliação do exame oral do IELTS.

Observando o Quadro 24 - resumo de critérios estruturais de análise 1, podemos notar que um problema continua a se repetir: a definição de critérios de avaliação voltados aos níveis C1 e C2 é variável e relativamente vaga. A delimitação clara do que torna, ou não, um falante de inglês como segunda língua, ou língua estrangeira, competente não é traçado em termos explícitos quanto à pronúncia (ex: domínio constante da vogal /æ/) ou quanto ao alcance gramatical (ex: utilização de inversões e inversões negativas).

À luz da ausência de pontuações que nomeiem critérios específicos, interpretamos os critérios estruturais mais marcantes aos níveis 7, 8 e 9, que ao nosso ver são, para o “alcance gramatical e precisão” - o controle gramatical, isto é, a quantidade e qualidade do erro cometido pelo candidato, e o domínio e competência em aplicar variadas formas sintática às próprias produções orais; para a pronúncia: domínio dos sons dos aspectos fonéticos, o conjunto de sons de uma língua (ROACH, 2009), e fonológicos, a relação característica que determinados sons pragmaticamente recebem por meio do uso de seus falantes (UNDERHILL, 1994), da língua em nível segmental, ou seja, os sons individuais, e supra segmental, isto é, do domínio dos aspectos rítmicos e tonais da fala (ROACH, 2009), com apenas ocasionais lapsos de pronúncia - cremos que isso seja referente, por exemplo, ao uso errôneo, pontual, de sons específicos da língua inglesa.⁵⁵

Quadro 24 - resumo de critérios estruturais de análise 1

Produção	Insumo das tarefas	Pronúncia - IELTS	Alcance gramatical e precisão - IELTS
As definições anteriores se aplicam a este nível, mas espera-se controle pleno das estruturas gramaticais.	As tarefas devem suscitar situações similares à anterior. Contudo, devem apresentar mais conteúdo abstrato possivelmente distante da realidade do candidato.	9 <ul style="list-style-type: none"> • usa uma gama completa de recursos de pronúncia com precisão e sutileza • mantém o uso flexível de recursos em toda a fala • é fácil de entender 	9 <ul style="list-style-type: none"> • usa uma gama completa de estruturas de forma natural e adequada • produz estruturas consistentemente precisas além das características de "deslizes" da fala do

⁵⁵ Todos os critérios referentes à fonética e fonologia do inglês foram definidos de acordo com Roach (2009) e Underhill (1994).

			falante nativo
As definições anteriores se aplicam a este nível, mas espera-se controle pleno das estruturas gramaticais.	As tarefas devem suscitar situações similares à anterior. Contudo, devem apresentar mais conteúdo abstrato possivelmente distante da realidade do candidato.	8 <ul style="list-style-type: none"> • usa uma ampla gama de recursos de pronúncia • mantém o uso flexível de recursos da fala conectada, com apenas lapsos ocasionais • é fácil de entender durante a fala; O sotaque da L1 tem efeito mínimo na inteligibilidade 	8 <ul style="list-style-type: none"> • usa uma ampla gama de estruturas de forma flexível • produz a maioria das sentenças livres de erros, apenas com inadequações muito ocasionais ou erros básicos / não sistemáticos
As definições anteriores se aplicam a este nível, com a adição de um critério que penaliza erros gramaticais que resultem em falhas de comunicação.	As tarefas devem suscitar situações similares à anterior. Contudo, devem apresentar mais conteúdo abstrato possivelmente distante da realidade do candidato.	7 <ul style="list-style-type: none"> • mostra todas as características positivas da Banda 6 e algumas, mas não todas, das características positivas da Banda 8 	7 <ul style="list-style-type: none"> • usa uma variedade de estruturas complexas com alguma flexibilidade • freqüentemente produz frases sem erros, embora alguns erros gramaticais persistam

Fonte: (VAN EK; TRIM, 1998a; 1998b; 2000; COUNCIL OF EUROPE, 2001; SEEDHOUSE; EGGBERT, 2004)

Utilizamos o trabalho de Roach (2009) e Underhill (1994) para caracterizar o que cremos ser a definição de “uso flexível com precisão e sutileza” das características da pronúncia. Segundo os autores, os aspectos mais marcantes da fala “natural”, a concatenação de elementos fonético-fonológicos (*features of connected speech*) de um determinado idioma são, além do domínio individual de cada um dos fonemas⁵⁶ da língua inglesa, como descrito no Quadro 24, a habilidade de conectar fonemas⁵⁷ de acordo com seu contexto de enunciação. Diferentemente de um contexto social, um contexto e de enunciação, de um ponto de vista meramente fonético-fonológico, é o momento da fala em que relações entre um dos fonemas das palavras utilizadas e outro se criam por meio de diferentes eventos.

⁵⁶ A menor unidade sonora de uma língua.

⁵⁷ É a menor unidade sonora com significado.

Na língua inglesa temos a *assimilação* (assimilation)⁵⁸, a *elisão* (*elision*)⁵⁹, a *catenação* (*catenation*)⁶⁰, *sons intrusivos* (*intrusive sounds*)⁶¹ e *formas fracas* (*weak forms*)⁶². Cada um desses eventos fonológicos ocorre em um contexto de fala diferente. A palavra “contexto”, neste parágrafo, refere-se somente aos contextos fonético-fonológicos gerados entre um fone e outro. Esses contextos podem, ou não, serem conectados a um ambiente social para uma investigação, por exemplo, de variação de registro ou ambiente de interação. No caso do exame oral do IELTS, cremos que a avaliação que os descritores fazem seja estrutural, voltando-se para a avaliação da precisão com que os sons são pronunciados, e não necessariamente seus contextos sociais. Ilustramos nossa opinião por meio de uma atividade hipotética.

Uma maneira de se inserir contexto social em um exame é por meio da atribuição de papéis pragmaticamente estabelecidos aos candidatos com base em uma situação interacional dramatizada (ex: dois amigos em um bar falando sobre futebol). Subsequentemente, outra situação de caráter interacional diverso, preferencialmente com outro tipo de registro relativo ao propósito da prova, poderia ser dramatizada para gerar amostras que favorecessem uma avaliação comparativa (ex: bar *versus* escritório) a fim de observar a capacidade do candidato de controlar determinadas características fonético-fonológicas pertencentes aos ambientes simulados, ou não.

Ao comparar tal proposição com as interações presentes no exame oral do IELTS por meio de suas tarefas (SEEDHOUSE et al, 2004), notamos que a maioria das tarefas apresenta diálogos predominantemente mediados pelo avaliador. Pouco pode ser dito em relação à multiplicidade de contextos socialmente relevantes ao candidato cujo propósito é acadêmico,

⁵⁸ Um fonema traz características para outro, podendo resultar em alterações de ponto de articulação, modo de articulação, e/ou vozeamento. A assimilação pode ser regressiva, ou progressiva. ex: fat boy. O /b/ torna-se um /p/.

⁵⁹ A elisão resulta da queda, ou omissão de um fonema similar em ponto e/ou modo de articulação, que se encontre próximo de si. Em “next day” o /t/ é elidido.

⁶⁰ A catenação, ou *linking*, é quando a junção de uma ou mais palavras ocorre em decorrência de padrões silábicos favoráveis para sua junção (consoante + vogal, vice-versa).

⁶¹ Sons intrusivos são sons que se inserem em contextos enunciativos específicos. Existem três sons intrusivos: o /j/, o /w/ e o /r/. O primeiro caso ocorre quando sons como /eɪ/, /i/, e /aɪ/ são sucedidos por uma vogal frontal, o segundo caso ocorre quando /ʊ/ and /oo/ são sucedidos por uma vogal arredondada recuada, e o terceiro ocorre sempre que uma palavra terminada em *schwa* é sucedida por vogal frontal não-arredondada.

⁶² Formas fracas são formas características dependentes do contexto de fala e se manifestam de acordo com a relação entre as palavras que expressão conteúdo presentes em uma sentença estão organizadas. Formas fracas, normalmente, resultam na redução de uma característica fonológica da forma forte da palavra reduzida, como em to /tu/. O /u/ passa a ser um *schwa* em sua forma fraca.

uma vez que as partes mais interacionalmente alinhadas a tal contexto sociocognitivo são as partes 2 e 3, que permitem, respectivamente, um turno longo de fala e uma conversa em moldes temáticos sociocognitivamente alinhados ao contexto acadêmico, mas sem, necessariamente, suscitar outras funções interacionais relativas aos múltiplos contextos fonético-fonológicos que podem permear tais ambientes (INOUE et al, 2018; VAN DIJK, 2012).

Ainda que tenhamos comentado que a grade de avaliação apresenta variada gama de referências linguísticas em sua composição, e que isso concede-a características de análise polivalentes, quando discorrendo sobre estrutura, há de se considerar que critérios como “deslizes da fala do falante nativo” pode possuir um contexto cujo fundamento social (VAN DIJK, 2012) dita não só a característica do erro, mas também a aceitabilidade do erro.

A base social que subjaz uma determinada interação enquadra, como já elucidado, uma série de questões pragmáticas que nos levam a questionar a aceitabilidade e a visibilidade de um erro, ou de um acerto. Tomemos como exemplo a discussão que Van Dijk (2012) traz à tona sobre Tony Blair, ex-primeiro ministro britânico. Vários exemplos sobre a liturgia do ambiente são postas à prova por meio da interação de seus integrantes. Turnos de fala, correções, roubos de turno de fala, permissão e proibição de enunciados complementares e/ou paralelos são exemplos referentes ao contexto social em que uma determinada interação se instaura. Em outras palavras os recursos linguísticos e, conseqüentemente, gramaticais, fonético-fonológicos, e interacionais, possuem caráter *egocêntrico* (VAN DIJK, 2012); ou seja, um contexto socialmente fundamentado possui suas próprias regras de “alcance gramatical e precisão”, que não parecem ter sido contempladas diretamente quanto ao contexto acadêmico na grade de avaliação do exame oral do IELTS.

Outrossim, é notável, então, sob perspectiva sociocognitivista, que as relações sociais que criamos, bem como o lugar que nascemos e as experiências que temos, influenciam a maneira como falamos, pensamos e agimos. Trata-se da formação ontogênica (VYGOTSKY, 1986; TOMASELLO, 2003) de um ser que, invariavelmente, é moldada de acordo com suas experiências em vida. Cremos que tal aspecto poderia ser contemplado também com nossas experiências com línguas estrangeiras, o que gera, naturalmente, diferentes noções de controle das características da fala e pronúncia. Um exemplo, em contexto de primeira língua, é o trazido por Gumperz (1972). O autor ocupou-se de investigar as variações presentes em

diferentes comunidades de fala e prática e constatou diferenças que variavam de acordo com a classe social, situação familiar e idade dos sujeitos.

Partindo agora de um segundo ponto de vista da realidade acadêmica para a investigação dos critérios estruturais, ao considerar que um ambiente acadêmico também é educacional, há um ponto que consideramos relevante a ser investigado no critério 7, que se difere dos critérios descritos em 8 e 9.

O critério 7 evidencia pontos analíticos de cunho comunicativo-estruturais necessários à investigação do desenvolvimento de um discente. Isto é, tal característica considera o caráter do erro de um falante avançado cujas habilidades ainda estão em desenvolvimento quando contrastado com falantes avaliados de acordo com os critérios elencados em 8 e 9. Uma vez que Weir (2005) aponta para a necessidade de um caráter educacional subjazendo um exame, e não somente uma relação entre propósito e exame, em termos de informação quanto ao desenvolvimento da proficiência do candidato, tal critério mostra-se, de certa forma, alinhado a fins acadêmicos, pois informa ao candidato em que ponto da grade ele está situado.

Assim sendo, o construto-alvo mobilizado nos parece ter mais características gerais do que específicas por desconsiderar, parcialmente, múltiplos ambientes que podem relacionar-se a fonética e fonologia, por exemplo. Trata-se, aparentemente, de um construto cuja base está calcada na competência comunicativa geral, a qual não abrange, necessariamente, as maneiras com que interações verbais e não-verbais se desenvolvem em ambientes acadêmicos. Isso não é dizer que uma avaliação de proficiência em língua estrangeira não deve testar competências gerais de comunicação. O que defendemos é: tanto em termos de pronúncia, quanto de gramática, as noções de “aceitabilidade”, “adequação” e “erro” são subjetivas e conectadas à contextos cujos modelos pragmáticos não podem ser definidos de maneira ambígua.

Dessa forma, cremos que os critérios “pronúncia” e “alcance gramatical e precisão” são avaliados de acordo com fundamentos comunicativo-estruturalistas na grade de avaliação do exame oral do IELTS. Como já descrito na investigação sobre “recursos lexicais”, entendemos que um mediador-avaliador que precisasse ter amplo domínio cultural para poder fazer uma avaliação de variados contextos de interação acabaria por ter grande carga cognitiva sobre si.

Quanto aos demais descritores, apesar de não haver ponto pacífico quanto aos critérios que resumem um falante de nível B2 em termos estruturais, o que podemos afirmar, pelo menos em relação aos níveis inferiores ao 7, é que há, principalmente em termos de proficiência geral, alinhamento com a literatura específica sobre competências sociocognitivas, dada a presença de critérios como “erro de caráter nativo”. No entanto, os níveis mais baixos possuem poucos descritores que nos ajudem a traçar mais considerações.

Não cremos que o exame oral do IELTS deva desenvolvê-los de maneira mais aprofundada dado que falantes com nível inferior a 5, de acordo com o Quadro 8, presente no capítulo metodológico, apresentariam grandes dificuldades para mobilizar, ou mesmo para organizar, recursos cognitivos referentes à prática interacional acadêmica em outra língua. Contudo, definições com distinções mais claras serviriam a uma melhor compreensão das relações que um investigador, ou mediador-avaliador, poderia traçar a partir da produção suscitada pelas tarefas e interpretação dos critérios da grade de avaliação.

Em resumo, apesar das considerações positivas sobre as análises passíveis de serem feitas com base na grade de avaliação do IELTS, a adequação para propósito acadêmicos encontra-se afetada e, portanto, consideramos sua validade de contexto desalinhada aos propósitos do exame, assim como parcialmente desalinhada em termos de propósito e insumo das tarefas. Nos Quadros 25 e 26 apresentamos os demais critérios entre as percepções que investigamos e a grade de avaliação do IELTS nas seções de “pronúncia” e “alcance gramatical e precisão”.

Quadro 25 - resumo de critérios estruturais de análise 2

Produção	Insumo das tarefas	Pronúncia IELTS	Alcance gramatical e precisão - IELTS
A produção gramatical esperada no nível B2 ainda é alvo de investigação na literatura especializada. As definições gramáticas referem-se à tentativa de estruturas complexas, maior concatenação de ideias e formas sintáticas atípicas.	As tarefas utilizadas para aferir candidatos do nível B2 são alvo de investigação, uma vez que não há especificação do que é considerado relevante ao nível. Destacam-se o uso de tarefas que tragam mobilizem o uso de formas verbais complexas, variedade	6 <ul style="list-style-type: none"> • usa uma variedade de recursos de pronúncia com controle misto • mostra algum uso eficaz de recursos, mas isso não é sustentado • geralmente pode ser totalmente compreendido, embora a pronúncia incorreta de palavras ou sons 	6 <ul style="list-style-type: none"> • usa uma mistura de estruturas simples e complexas, mas com flexibilidade limitada • pode cometer erros frequentes com estruturas complexas, embora raramente causem problemas de compreensão

	sintática e modais de especulação.	individuais reduza a clareza às vezes	
A produção esperada no nível B1 alto se diferencia B1 baixo por avaliar a tentativa de uso de estruturas pertencentes ao nível seguinte.	As tarefas utilizadas para aferir o desenvolvimento de candidatos desse nível costuma avaliar se as formas básicas utilizadas pelos candidatos organizam-se de maneira controlada, isto é: descrever e expor itens da realidade imediata do candidato sem erros gramaticais básicos com tentativas de estruturas do nível seguinte.	5 <ul style="list-style-type: none"> • mostra todas as características positivas da Banda 4 e algumas, mas não todas, das características positivas da Banda 6 	5 <ul style="list-style-type: none"> • produz frases básicas com precisão razoável • usa uma gama limitada de estruturas mais complexas, mas geralmente contém erros e podem causar alguns problemas de compreensão

Fonte: (VAN EK; TRIM, 1998a; 1998b; 2000; COUNCIL OF EUROPE, 2001; SEEDHOUSE; EGGBERT, 2004)

Quadro 26 - resumo de critérios estruturais de análise 3

Produção	Insumo das tarefas	Pronúncia IELTS	Alcance gramatical e precisão - IELTS
A produção esperada no nível B1 não se diferencia tanto do nível A2. O que espera-se é maior controle estrutural das formas gramaticais básicas da língua-alvo.	As tarefas utilizadas para aferir o desenvolvimento de candidatos desse nível costuma avaliar se as formas básicas utilizadas pelos candidatos organizam-se de maneira controlada, isto é: descrever e expor itens da realidade imediata do candidato sem erros gramaticais básicos.	4 <ul style="list-style-type: none"> • usa uma gama limitada de recursos de pronúncia • tentativas de controlar recursos, mas lapsos são frequentes • erros de pronúncia são frequentes e causam alguma dificuldade para o ouvinte 	4 <ul style="list-style-type: none"> • produz formas de frases básicas e algumas frases simples corretas, mas estruturas subordinadas são raras • erros são frequentes e podem levar a mal-entendidos
A produção esperada de candidatos desse nível é somente inerente à formas gramaticais simples. Um candidato pode ser mostrar-se preciso em estruturas simples e, ao mesmo tempo, se arriscar mais para a produção de sentenças contendo ideias mais complexas.	As tarefas utilizadas para aferir o desenvolvimento de candidatos desse nível não costuma elencar funções fora do escopo funcional básico: descrever e expor itens da realidade imediata do candidato.	3 <ul style="list-style-type: none"> • mostra algumas das características da Banda 2 e algumas, mas não todas, as características positivas da Banda 4 	3 <ul style="list-style-type: none"> • tenta formas de frases básicas, mas com sucesso limitado, ou depende de declarações aparentemente memorizadas • comete vários erros, exceto em expressões memorizadas
		2 ⁶³	2

⁶³ Não há descritores presentes para os níveis 0 e 1 nos quesitos “recurso lexical”, “alcance gramatical” e “pronúncia”.

		• A fala muitas vezes é ininteligível	• não pode produzir formas de frases básicas
--	--	---------------------------------------	--

Fonte: (VAN EK; TRIM, 1998a; 1998b; 2000; COUNCIL OF EUROPE, 2001; SEEDHOUSE; EGGBERT, 2004)

5.15.3 Análise de aspectos funcionais do insumo e da grade de avaliação

Os aspectos funcionais do insumo já foram parcialmente discutidos ao longo deste capítulo de análise. Para Galaczi e French (2011), os aspectos funcionais do insumo e da grade de avaliação da língua-alvo são centrais para definir o que os candidatos são capazes de fazer por meio da língua. Alinhamo-nos com a opinião das autoras, dado que uma investigação de cunho sociocognitivo deve focar a maneira como as funções discursivas manifestam-se ao longo das interações contextualmente situadas. Índícios de mobilização interacionais obtidos por meio de funções podem ser, também, indicativos de capacidades cognitivas relativas a aspectos gramaticais e lexicais específicos (GALACZI; FRENCH, 2011).

Assim sendo, as interações, segundo o QECR (COUNCIL OF EUROPE, 2001), podem ser organizadas em duas categorias: a categoria macrofuncional e a categoria microfuncional. As macrofunções relacionam-se às características da informação presentes em uma interação. Um ponto de partida para melhor entender tal critério é a definição de conversa-factual e avaliativa de Bygate (1987). As microfunções (COUNCIL OF EUROPE, 2001) são pertinentes a atividades interacionais subordinadas às rotinas interacionais que as enquadram, aliando-se às funções discursivas descritas em Brown e Yule (1983). Da mesma forma, O'Sullivan (2002) e suas categorias interacionais também são úteis por relacionarem-se diretamente com os critérios previamente citados.

As definições do QECR (COUNCIL OF EUROPE, 2001) também alinham-se às definições de macro e microcontexto, assim como à definição de planejamento de Van Dijk (2012). Em Van Dijk (2012), microcontextos são modelos de interação sociais momentâneas que podem ser representações de situações históricas (macrocontextos). O planejamento vai diferir de acordo com os macro e microcontextos instaurados em um determinado ambiente social em decorrência dos esquemas mentais e pragmáticos resultantes de cada uma de suas

realidades sociohistórico-culturais. Em outras palavras, o contexto interacional dita a simulação das tarefas e dos padrões interacionais que os interactantes devem adequar-se, além de mobilizar, na mente dos candidatos, esquemas de planejamento prévio da pragmaticamente convencionados e co-construídos. Um exemplo seria uma conversa entre amigos que resulta em conflito e evoca microfunções contextuais, como a de reparação da interação (GOFFMAN, 1967). Assegurar validade de contexto é assegurar a prevalência de tais características interacionais.

É possível inferir, então, que a definição dos possíveis aspectos funcionais de uma tarefa nos direciona para horizontes sociocognitivos da confecção de exames. Ora, não basta descrever, comparar, concordar e/ou discordar. Tal como descrito nas análises dos critérios de “alcance gramatical e precisão”, “pronúncia” e “léxico”, assim como em Weir (2005), os interactantes devem ter em mente *quando, se e como* realizar tais interações com base em seus papéis contextualmente definidos. Logo, a investigação dos aspectos funcionais de um exame oral depende da definição dos macro e microcontextos que translacionam e, ao mesmo tempo, são translacionados pela interação humana. O’Sullivan (2002) elenca categorias funcionais que podem nos auxiliar a investigar as macro e microfunções de uma tarefa. A primeira é a categoria informacional, que atribui o tipo de informação e função a serem produzidas pelo candidato. A segunda é a interacional, atribuindo funções específicas aos contextos suscitados por meio do enunciado das tarefas. A terceira, a regulação das funções interacionais, é relativa à segunda, mas delega mais responsabilidade aos interactantes trazendo à tona, por exemplo, a necessidade de negociação de significados.

Apresentamos, mais a frente, três quadros que interrelacionam demandas funcionais das tarefas e grade de avaliação de acordo as categorias conversacionais de Bygate (1987), funções discursivas de Brown e Yule (1983), categorias macro e microfuncionais do QECR (COUNCIL OF EUROPE, 2001) e categorias interacionais de O’Sullivan (2002). Os Quadros 27, 28 e 29, são utilizados de acordo com nossa interpretação do que cada parte do exame mobiliza quanto ao construto-alvo em relação a seu propósito; isto é, o alinhamento entre as funções das tarefas, grade de avaliação e validade de contexto. O uso das funções listadas, tal como elencado nas análises anteriores, também depende do nível de proficiência dos candidatos. Logo, o uso “preciso” e “satisfatório” da função discursiva e interacional necessária para realizar a tarefa depende da proficiência do candidato, pois, pela função

também serão mobilizados recursos cognitivos referentes ao conhecimento prévio do candidato. Assim, podemos dizer que, quanto maior for a demanda interacional de uma tarefa, maior será a mobilização de recursos sociocognitivos, o que pode influir no desempenho de candidatos com menos esquemas mentais pré-definidos no idioma-alvo

Assim como nos aspectos gramaticais e lexicais, o material referente às funções interacionais dos níveis C1 e C2 ainda são alvo de estudo e discussão. Portanto, utilizaremos, tal como nas outras seções, nosso referencial prévio (VAN EK; TRIM, 1998a; 1998b; 2001; COUNCIL OF EUROPE, 2001) para desenvolver a investigação e tecer a costura de nossa análise.

De um ponto de vista em que estrutura e função se relacionam, diferentes níveis de proficiência podem ser analisados a partir da quantidade de funções e modelos pragmaticamente funcionais suscitados em relação ao contexto simulado. Por exemplo, em uma tarefa de opinião, pode-se diferenciar um candidato de nível mais baixo de um de nível mais alto pela produção; uma resposta como “*I like this.*” (eu gosto disso) distancia-se muito de uma como “*I think I would choose this because...*” (eu acredito que escolheria isso porque...). Claramente, o candidato que produzisse a segunda amostra seria considerado mais hábil em termos de uso da língua do que o primeiro.

Ademais, a previsibilidade da tarefa, isto é, o quão concreto ou abstrato o assunto evocado é, relaciona-se diretamente à proficiência do candidato. Assuntos muito próximos da realidade comum ao dia-a-dia são considerados assuntos de baixa previsibilidade, enquanto, quanto mais distantes, mais imprevisível se tornam. A Figura 8 exemplifica a ideia exposta somente em algumas instâncias, pois ela não define o nível de complexidade, ou temática, que assuntos podem apresentar em termos de dificuldade e/ou previsibilidade de maneira definitiva.

**Quadro 27 - resumo de funções
informacionais em ordem de complexidade**

<p>Expor informações pessoais</p> <ul style="list-style-type: none"> - presente - passado - futuro
Expressar opiniões
Elaborar
Justificar opiniões
Comparar
Especular
Dramatizar
<p>Descrever</p> <ul style="list-style-type: none"> - Sequências de eventos - Eventos
Organizar/listar
Sugerir
Expressar preferências/emoções

Fonte: (COUNCIL OF EUROPE, 2001; O'SULLIVAN, 2002; TAYLOR, 2011)

**Quadro 28 - resumo de funções
interacionais em ordem de complexidade**

Concordar
Discordar
Modificar/comentar a informação dada
Pedir opinião
Persuadir
Pedir por informação
Reparar a comunicação
<ul style="list-style-type: none"> - Negociar significado - Checar significado - Demonstrar compreensão - Procurar opiniões comuns aos interactantes - Pedir esclarecimentos - Corrigir frases <p>responder/esclarecer dúvidas do outro interactante</p>

Fonte: (COUNCIL OF EUROPE, 2001; O'SULLIVAN, 2002; TAYLOR, 2011)

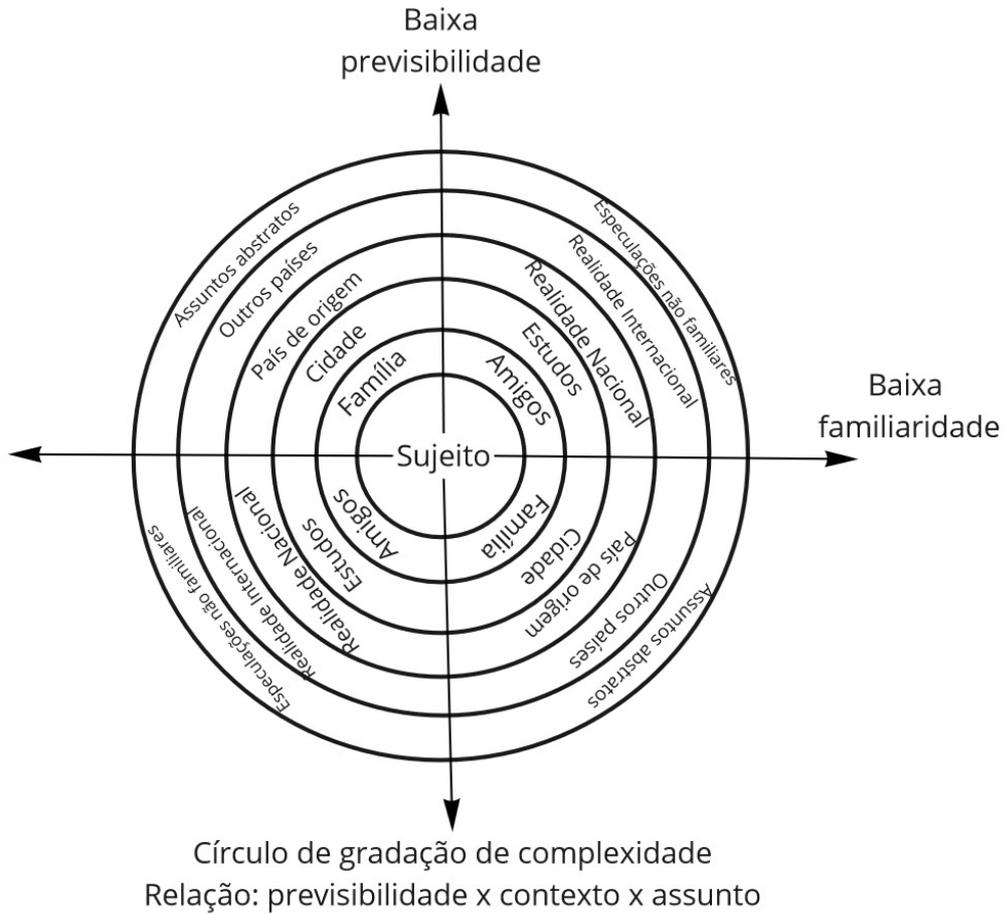
Quadro 29 - resumo de gerenciamento de funções interacionais

Iniciar a interação
Mudar o tipo e assunto da interação
Demonstrar interesse recíproco na interação
Decidir com base nas opções dadas na interação

Fonte: (COUNCIL OF EUROPE, 2001; O'SULLIVAN, 2002; TAYLOR, 2011)

Resta-nos organizar a maneira como as funções tendem a ser distribuídas por nível de proficiência em contexto geral para poder analisá-las no contexto de interação simulado no IELTS, junto de seus propósitos.

Figura 8 - gradação de complexidade temática



Fonte: Elaboração própria.

Quadro 30 - resumo das macro e microfunções e previsibilidade da produção organizadas por nível de proficiência

Nível	Macrofunções seguidas de microfunções características	Microfunções interacionais	Microfunções de regulação interacional	Previsibilidade das tarefas em relação a contextos sociais
A1/A2/ B1	Em ordem de complexidade Conversa-factual	Não obrigatórias	Não obrigatórias	Alta

	<ul style="list-style-type: none"> - Descrever - Organizar - Sugerir 			
B2	<p>Em ordem de complexidade</p> <p>Conversa-factual e conversa-avaliativa</p> <ul style="list-style-type: none"> - Descrever - Organizar - Sugerir - Especular - Dramatizar <p>Sequenciar/organizar eventos</p>	<ul style="list-style-type: none"> - Interagir socialmente; - Trocar informações <ul style="list-style-type: none"> - Persuadir - Expressar preferências - Expressar emoções 	Contextual	Média
C1/C2	<p>Em ordem de complexidade</p> <p>Conversa-factual e conversa-avaliativa, com prevalência da avaliativa</p> <ul style="list-style-type: none"> - Descrever - Organizar - Sugerir <p>Especular com mais profundidade</p> <ul style="list-style-type: none"> - Dramatizar <p>Sequenciar/organizar eventos com mais elementos</p>	<ul style="list-style-type: none"> - Interagir socialmente; - Trocar informações <ul style="list-style-type: none"> - Persuadir - Expressar preferências - Expressar emoções - Construir opiniões abstratas 	Contextual	Baixa

Fonte: (BYGATE; 1987; BROWN; YULE, 1983; COUNCIL OF EUROPE, 2001; O’SULLIVAN, 2002; TAYLOR, 2011)

A seguir, apresentamos análise do nível de complexidade do conteúdo das tarefas, suas macro e microfunções e assim como os tipos de gerenciamento de interação presentes no exame oral do IELTS. As análises são apresentadas por meio de oito figuras diferentes. Reiteramos que somente a parte 1 apresentava diferentes temas em suas perguntas em nossas amostras (SEEDHOUSE et al, 2004), enquanto as partes 2 e 3 do exame possuíam o mesmo formato, número de perguntas e temas. As macrofunções (COUNCIL OF EUROPE, 2001) foram utilizadas em consonância às rotinas interacionais de Bygate (1987), enquanto a análise das microfunções (COUNCIL OF EUROPE, 2001) foi aliada às categorizações de Brown e Yule (1983) e utilizadas de acordo com as tarefas. Outros aspectos das microfunções das tarefas foram alocados na seção de regulação interacional. As questões utilizadas em nossos

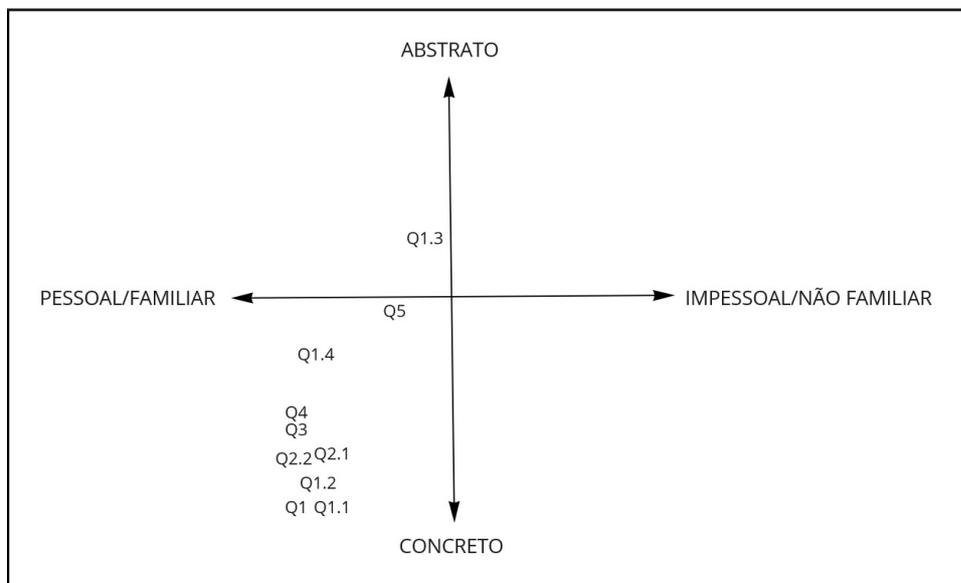
excertos de entrevista foram organizadas em códigos para facilitar sua visibilidade e referência nas seções posteriores do texto. As questões de cada uma das partes serão referenciadas por seus respectivos códigos. As interpretações que fizemos refletem nossa interpretação do tema.

Quadro 31 - Análise da parte 1 da entrevista 1 do exame oral do IELTS

<p>Macrofunções Conversa-factual levemente avaliativa</p> <p>Microfunções Descrição Descrição Opinião Opinião Exposição/Especulação Opinião Opinião Comparação/Especulação Opinião/Exposição Opinião Exposição</p> <p>Regulação interacional Mediador: iniciar e manter o diálogo Candidato: responder às perguntas</p>	<p>Parte 1 - Entrevista 1</p> <p>Q1 - Can you tell me where you are from? Q2 - Can we talk about where you live? Could you describe where you live? Q3 - What do you like about living there? Q4 - Are there things you don't like about it? Q5 - How could you improve the city?</p> <p>Let's move on to the topic of food and restaurants</p> <p>Q1.1 - What kind of food do you like to eat? Q1.2 - Is there any food you don't like? Q1.3 - What are some of the advantages and disadvantages of eating in a restaurant? Q1.4 - What's the good thing about eating in a restaurant?</p> <p>Let's talk about films?</p> <p>Q2.1 - Do you enjoy watching films? Q2.2 - How often do you watch films?</p>
--	--

Fonte: (BYGATE; 1987; BROWN; YULE, 1983; COUNCIL OF EUROPE, 2001; O’SULLIVAN, 2002; TAYLOR, 2011)

Figura 9 - Conteúdo das tarefas da parte 1 entrevista 1 do exame oral do IELTS



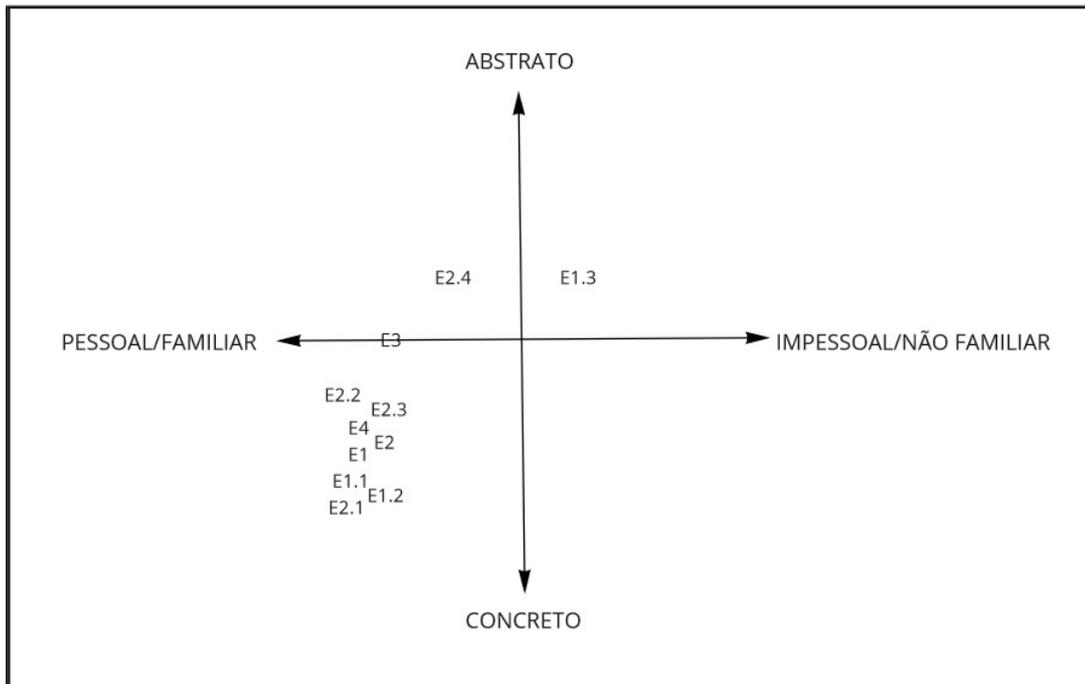
Fonte: (GALACZI; FFRENCH, 2011; COUNCIL OF EUROPE, 2001)

Quadro 32 - Análise da parte 1 da entrevista 2 do exame oral do IELTS

<p>Macrofunções Conversa-factual levemente avaliativa</p> <p>Microfunções Descrição Opinião Opinião Exposição/Especulação</p> <p>Exposição Exposição Especulação</p> <p>Descrição Argumentação/Persuasão. Descrição Especulação</p> <p>Regulação interacional Mediador: iniciar e manter o diálogo Candidato: responder às perguntas</p>	<p>Parte 1 - Entrevista 2</p> <p>E1 - Tell me about your studies E2 - What do you like most about your studies? E3 - Are there any things you don't like about your studies? E4 - What qualifications or certificates do you hope to get?</p> <p>Let's talk about activities you enjoy in your free time</p> <p>E1.1 - when do you have free time? E1.2 - What free time activities are most popular where you live? E1.3 - How important is free time in people's lives?</p> <p>Can we talk about your childhood?</p> <p>E2.1 - Where did you grow up? E2.2 - Was it a good place for children? why? E2.3 - Where did you usually play? E2.4 - Do you think childhood is different today than when you were a child?</p>
---	---

Fonte: (BYGATE; 1987; BROWN; YULE, 1983; COUNCIL OF EUROPE, 2001; O’SULLIVAN, 2002; TAYLOR, 2011)

Figura 10 - Análise do conteúdo das tarefas da parte 1 entrevista 2 do exame oral do IELTS



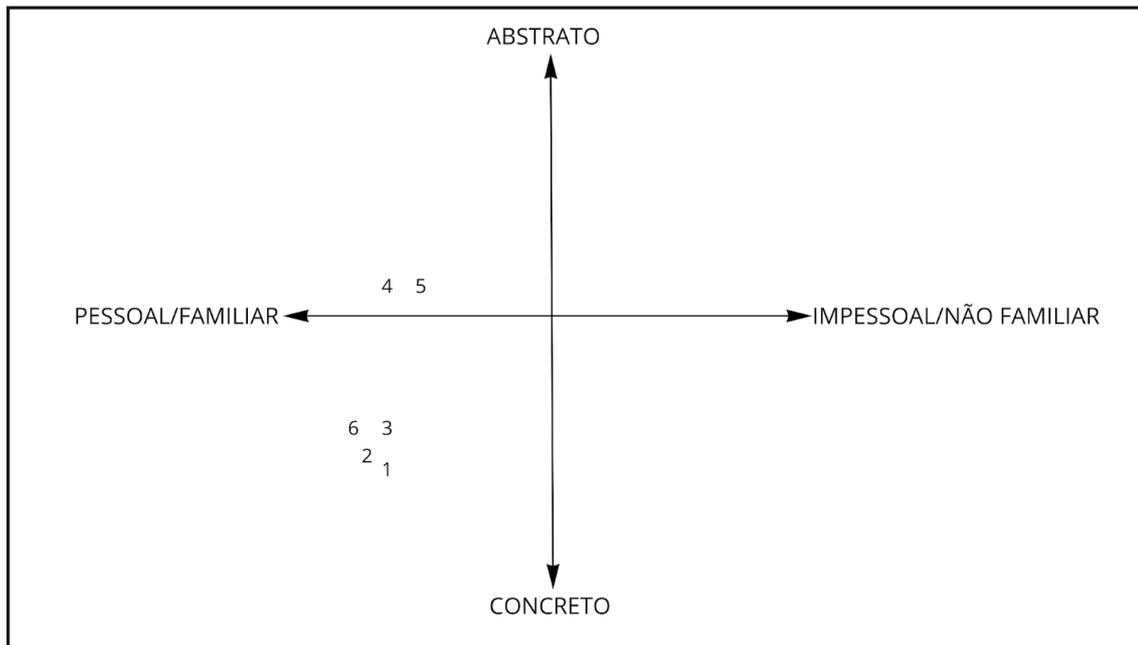
Fonte: (GALACZI; FFRENCH, 2011; COUNCIL OF EUROPE, 2001)

Quadro 33 - Análise da parte 2 do exame oral do IELTS

<p>Macrofunções Conversa-factual e conversa-avaliativa</p> <p>Microfunções Organizar/Sequenciar Exposição Exposição/Descrição Descrição Argumentação/Persuasão Especulação</p> <p>Exposição</p> <p>Regulação interacional Mediador: iniciar o diálogo e instruir o candidato Candidato: produzir turno longo de fala coerentemente</p>	<p>Parte 2 - Turno longo de fala</p> <p>Now let's move on to the next part and I am going to give you a topic, and I'd like you to talk about it for one or two minutes. Before you talk you will have one minute to think about what you are going to say. And you can make some notes if you wish. All right? Do you understand?</p> <p>Here's some paper for making notes and here's your topic. I'd like you to describe a trip that you once went on.</p> <p>(written on task card)</p> <p>1 - Where you went 2 - Why you went there 3 - How you traveled 4 - What you enjoyed most 5 - If you would repeat this trip and why</p> <p>6 -Do you generally enjoy traveling?.</p>
--	---

Fonte:(BYGATE; 1987; BROWN; YULE, 1983; COUNCIL OF EUROPE, 2001; O'SULLIVAN, 2002; TAYLOR, 2011).

Figura 11 - Análise do conteúdo das tarefas da parte 2 do exame oral do IELTS



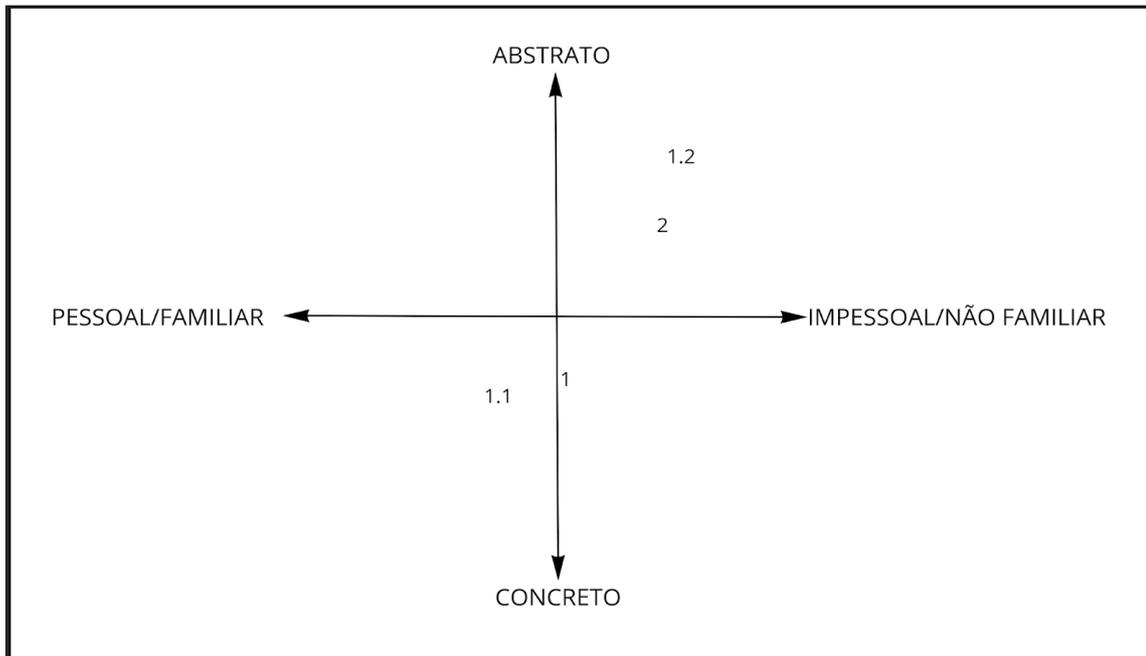
Fonte: (GALACZI; FFRENCH, 2011; COUNCIL OF EUROPE, 2001)

Quadro 34 - Análise da parte 3 do exame oral do IELTS

<p style="text-align: center;">Macrofunções</p> <p>Conversa-factual com mais elementos avaliativos</p> <p style="text-align: center;">Microfunções</p> <p>Comparação Exposição</p> <p>Comparação Especulação</p> <p>Descrição Exposição</p> <p>Especulação Argumentação/ Persuasão</p> <p style="text-align: center;">Regulação interacional</p> <p>Mediador: iniciar e assegurar continuidade da interação.</p> <p>Candidato: produzir amostra coerente com a natureza da interação</p>	<p>Parte 3 - Discussão aberta</p> <p>We've been talking about a trip that you went on and I'd like to discuss with you one or two more general questions related to this..</p> <p>Tópico 1 - Let's think about travel and transport.</p> <p>1 - What's the most popular way to travel long distance in your country?</p> <p>2 - How do you think we're going to travel in the future?</p> <p>Tópico 2 - Let's talk about public and private transportation.</p> <p>1.1 - Can you describe the public transport system in your country?</p> <p>1.2 Can you speculate on any measures that will be taken to reduce pollution in the future?</p>
---	---

Fonte: (BYGATE; 1987; BROWN; YULE, 1983; COUNCIL OF EUROPE, 2001; O'SULLIVAN, 2002; TAYLOR, 2011)

Figura 12 - Análise do conteúdo das tarefas da parte 3 do exame oral do IELTS



Fonte: (GALACZI; FFRENCH, 2011; COUNCIL OF EUROPE, 2001)

Se retomarmos o Quadro 10 que adaptamos de Ducasse e Brown (2011), podemos ter uma noção das relações existentes entre as funções mobilizadas nas tarefas investigadas em nossa dissertação e as funções discursivas mobilizadas em contexto acadêmico. Para tal, organizamos um segundo quadro comparativo, o Quadro 35. Em nossa análise, consideramos a possibilidade de mobilização de funções análogas às pesquisadas por Ducasse e Brown (2011) e as classificamos como presente, ausente e possível.

A investigação de Ducasse e Brown (2011) revela que somente seis das dezesseis funções investigadas foram mobilizadas durante as tarefas do exame oral do IELTS examinadas em suas amostras. Como partimos de um corpus de amostra diferente e de menor volume que se resume a dois exames documentados, não partimos, como as autoras, da produção. Partimos, ao contrário, das possíveis situações de mobilização de construtos de cunho acadêmico.

Assim, classificamos nossa análise com base em hipóteses criadas a partir da investigação das tarefas e das situações interacionais e funcionais possivelmente arquetípicas do contexto acadêmico. Consideramos, então, mais do que a tarefa em si em nossa análise, investigando até o que consideramos que essa poderia mobilizar. Não temos intenção de refutar, com nossa análise, a pesquisa de Ducasse e Brown (2011), afinal, a pesquisa das autoras foi calcada na produção e investigação dos dados produzidos pelos candidatos do exame, o que gera uma evidência concreta em favor de seu argumento. Nosso objetivo é, tão-somente, investigar a maneira com que as tarefas mobilizam construtos que sirvam para propósitos acadêmicos.

A análise foi conduzida a partir das funções interacionais que tomamos como possíveis. As classificadas como “presente” ou “ausente” são ponto pacífico em nossa concepção e também na de Ducasse e Brown (2011). Dessa forma, começamos a investigação pelas opções marcadas como “possíveis” a partir das tarefas investigadas. Vale ressaltar, novamente, que consideramos os pressupostos de Luoma (2004), os quais concebem a generalização das funções discursivas de um exame sob alçada temática variável como prática válida, pois elas nos permitem interpretar mais aspectos sociocognitivos das questões.

Quadro 35 - Análise das funções presentes em Ducasse e Brown (2011)

Análise	Funções análogas nas tarefas investigadas	Tipo de função	Funções presentes em Ducasse e Brown (2011)	Lista
Presente	Organizar/Sequenciar	Informação 1	Prover informações, ideias ou opiniões	X
Presente	Expor	Informação 2	Defender as próprias ideias ou opiniões	X
Possível	Argumentar/Persuadir	Informação 3	Elaborar as próprias ideias ou opiniões	
Possível	Opinar	Informação 4	Sugerir	
Possível	Especular			
Ausente	Comparar	Interação 1	Desafiar ideias de outro falante	
Ausente	Argumentar/Persuadir	Interação 2	Justificar/prover suporte a outro	
Possível	Limitada	Interação 3	Concordar/Discordar de outro falante	X
Possível	Expandir/Modificar ou ao reproduzir	Interação 4	Quantificar/modificar de outro falante	X
Possível	Argumentar/Persuadir	Interação 5	Persuadir outro falante	
Possível	Limitada	Interação 6	Pedir informação, ideias ou opiniões de outro falante	X
Possível	Limitada	Interação 7	Solicitar elaboração/justificativas de outro falante	X
Possível	Limitada	Interação 8	Desenvolver/modificar a ideia de outro falante	
Possível	Limitada	Interação 9	Negociar significado	X
Ausente	Limitada	Gerenciamento 1	Iniciar interação	
Ausente	Limitada	Gerenciamento 2	Mudar de assunto	
Possível	Limitada	Gerenciamento 3	Concluir um argumento/decidir	

Fonte: Elaboração própria; (DUCASSE; BROWN, 2011, p.12)

Ducasse e Brown (2011) não constataram a presença das informações 3 e 4 na produção de seus colaboradores. Contudo, cremos que a produção de conteúdo relativo a essas informações seja possível, particularmente nas tarefas E2.4 e E1.3, presentes na Figura 10.

As questões E2.4 e E1.3, cujas traduções aproximadas são “Você acha que a infância hoje é diferente de quando você era criança?” e “Qual a importância do tempo livre na vida das pessoas?”, como impossíveis de receberem réplica no contexto simulado no exame oral do IELTS, estaríamos, por exemplo, negando a capacidade heurística de transformação da informação, habilidade inerentemente pertencente ao ser humano. Tomasello (2003) analisa-a, principalmente, quanto ao uso de instrumentos (materiais ou simbólicos), em que novas funções são criadas a partir de reinterpretações de seu significado.

Ressaltamos que não desconsideramos a possibilidade de candidatos simplesmente se recusarem a produzir silogismos para construir argumentos, como no caso histórico das primeiras pesquisas cognitivas citadas em Cole (1996). Uma breve paráfrase silogístico-especulativa que responderia e atenderia os critérios de Ducasse e Brown (2011) na

pergunta E2.4 seria, nesses moldes, algo como “Crianças antigamente saíam e brincavam na rua. Hoje, elas não o fazem, por isso, são gordas.”, em que uma opinião seria dada com base na experiência e percepção de mundo do candidato; da mesma forma, candidatos que não se sentissem confortáveis com uma especulação desse cunho, mas fossem proficientes, poderiam produzir menos.

A elaboração de ideias, o critério 3 do Quadro 35, é fundamental ao contexto acadêmico. No caso da questão E2.4, notamos a possível presença de habilidades inerentes à criação de hipóteses. Podemos inferir que uma das razões para os colaboradores de Ducasse e Brown (2011) não terem desenvolvido suas respostas de maneira satisfatória pode relacionar-se à falta de conexão com o assunto apresentado nas tarefas de seus exames, o que, possivelmente, pode desencorajar candidatos de responder a uma tarefa.

Outro motivo para um desenvolvimento incompleto do critério 3 do Quadro 10 pode estar relacionado à estrutura da questão que, por mais que permita a criação de hipóteses e elaboração de ideias, está estruturada como uma pergunta de sim ou não - fator decisivamente importante para seu processamento cognitivo.

Por fim, há de se considerar o que “elaborar” pode significar em contexto para que possamos investigar termos como “adequado”, “suficiente” ou “flexível”. No caso estabelecido por Ducasse e Brown (2011), o significado de “elaborar” é elucidado como “resposta [aparentemente] insuficiente, [com participação de] outros alunos [que também] elaboravam [ideias] por turnos iniciados por si mesmos, ou com o incentivo do professor à classe.” (DUCASSE; BROWN, 2011, p.15)⁶⁴. Consideramos, então, o critério “suficiente” em Ducasse e Brown (2011) tão difícil de ser delineado quanto critérios como “precisão”, ou “flexível”, presentes nas seções anteriores.

Se tratando de um exame oral cujo propósito da pergunta não elenca critérios a serem mencionados na resposta da tarefa, como na parte 2, em que comandos são posicionados para gerar um melhor entendimento do que “suficiente” viria a ser dentro da tarefa, a interpretação de tal critério fica à mercê do mediador-avaliador, assim como do candidato que a responde, o que potencialmente gera duplicidade quanto ao conceito de “suficiência”. Em outras palavras, a percepção do que vem a ser “suficiente” quanto à elaboração das ideias se tornam alvo de

⁶⁴ No original - When a response appeared to be insufficient, other students extended and elaborated it with student-initiated turns, or the lecturer elicited more from the group or the class.

subjetividade tanto para produção, quanto para avaliação. Nesse caso, os recursos dêiticos presentes no discurso de um sujeito (VAN DIJK, 2021) parecem ter significativa influência na produção do candidato porque ditam critérios inerentes à, por exemplo, extensão da resposta considerada aceitável, ou resposta suficientemente polida à interação.

Na questão E1.3, “Qual a importância do tempo livre na vida das pessoas?”, a mesma questão dêitica se aplica, mas, desta vez, voltada ao termo “sugerir”, mencionado no critério 4 do Quadro 35, uma vez que precisamos definir o que as autoras Ducasse e Brown (2011) entendem por esse termo.

Os tópicos foram negociados e sugestões foram feitas e apoiadas sobre como dividir os tópicos entre os membros da classe (...) Os alunos também fizeram sugestões sobre maneiras de reformular o que foi dito para a resposta escrita. (...) Outras interações exclusivamente discentes exigiam que os alunos compartilhassem informações que haviam pesquisado individualmente e que, portanto, eram conhecidas apenas por eles mesmos (por exemplo, a atividade de encenação). Aqui, deram informações, sugestões e responderam aos pedidos de apoio e elaboração. (DUCASSE; BROWN, 2011, p.15-16)⁶⁵

“Sugerir” parece ir além da descrição que as autoras trouxeram em sua análise presente no Quadro 35. De acordo com a passagem transcrita anteriormente, a ação de “sugerir” parece estar mais voltada ao ato de “interferir”, “adicionar” ou, quiçá, “criticar”.

No contexto da questão utilizada para exemplificar a ideia de “sugestão” no exame oral do IELTS, com base em nossa amostra, inferimos que seria possível para um candidato produzir enunciados alinhados à função da questão e, conseqüentemente, ao contexto acadêmico. Por exemplo, “Eu acho que as pessoas deveriam ter tempo livre para que possam meditar. Meditar faz bem à mente.”. Apesar de curta, a resposta possui uma opinião pessoal e uma justificativa, que, em si, carrega uma sugestão que direcionaria a importância e uso do que pode ser feito com tempo livre. A sugestão seria contemplada na brevidade da resposta. Acreditamos que ela possa ser considerada insatisfatória por não apresentar interação com uma informação cuja origem não seja outro falante, ou seja, de baixa natureza especulativa.

⁶⁵ No original - The topics were negotiated and suggestions were made and supported as to how to divide the topics among the members of the class (...) Students would also make suggestions about ways of rephrasing what had been said for the written answer. (...) Other student-only interactions required students to share information that they had individually researched and which was, therefore, known only to them (eg, the role-play activity). Here, they provided information and made suggestions, and responded to requests for support and elaboration. p.15-16

Mas, ainda assim, em termos macrocontextuais, poderia carregar as características de uma sugestão.

Partindo para as interações 3 e 4, de acordo com as considerações de Ducasse e Brown (2011) “concordar”, “discordar”, “quantificar” e “modificar” faziam parte da maneira com que “os alunos contribuíram para a discussão apoiando-se, concordando ou discordando”⁶⁶ (2011, p.16). Considerando que as produções investigadas por Ducasse e Brown (2011) partem de outras tarefas com, possivelmente, divergências em relação às funções elencadas em nossa amostra, cremos que as funções de interação 3 e 4 podem ter sido trazidas à tona com mais facilidade. Em nosso caso, as questões Q3, “Do que você gosta quanto a morar lá?”, e Q4, “Há coisas que você não gosta?”, presentes no Quadro 31, parecem ser as mais alinhadas à mobilização das interações de cunho acadêmico.

Apesar das perguntas não apresentarem, ao menos de maneira explícita, as palavras “concordar” e “discordar”, há de se considerar que um candidato pode aludir à opinião de terceiros para construir a sua, além de mencioná-los, gerando assim uma condição de alinhamento ou desalinhamento quanto ao tema da pergunta em sua produção. Caso um candidato respondesse a pergunta Q3 utilizando uma resposta como “Há quem não goste, mas eu gosto desse aspecto...”, há um valor implícito, novamente, dêitico (VAN DIJK, 2012), em que o alinhamento está expresso de forma ideológica, tão-somente, mas que não deixa de considerar um meta-discurso interacional instaurado nos esquemas mentais sociocognitivos pré-existentes desse usuário da língua. O ato de concordar, ou discordar, ocorre na instância mental, e é expresso de maneira referencial e abstrata, sem a especificação do sujeito que discorda.

Os atos de “expandir”, ou “modificar”, podem partir de um referencial cognitivo suscitado da mesma forma. O processo, contudo, seria ligeiramente diferente. A alusão que fazemos a Van Dijk é calcada em sua diferenciação, e ao mesmo tempo, aproximação de contexto e texto (VAN DIJK, 2012). Acreditamos que o discurso criado por um candidato para responder à pergunta Q4 pode, por exemplo, ser co-construído com base em textos e/ou discursos de outros falantes. Apesar de não se tratar de uma expansão explícita da opinião/discurso de um interactante na interação, tomando a acepção de Van Dijk (2012) sobre texto e contexto, que considera inferências de propriedades da formação de um discurso

⁶⁶ No original: “students contributed to the discussion by supporting, agreeing, or disagreeing.”

plausível à identificação de seus contextos de formação, podemos considerá-la aceitável, tal como as autoras (DUCASSE; BROWN, 2011). Ademais, aplicar-se-ia os conceitos previamente mencionados de Vähäpääsi (1982), de reprodução e organização da informação. As conclusões referentes à Q4 também se aplicariam à interação 8, uma vez que ela se calca em pressupostos similares, ainda que não tenha sido constatada pelas autoras (DUCASSE; BROWN, 2011).

Escolhemos desenvolver as interações 3 e 4 porque não conseguimos notá-las explicitamente nas perguntas de nossa amostra. Ademais, Ducasse e Brown (2011) demonstram que, por mais que tenha havido expansão e modificação de outros discursos na fala de seus colaboradores, foi somente em decorrência da provocação do mediador-avaliador, com candidatos de alto nível de proficiência. O discurso expandido foi o deles próprios. Consideramos que seria necessário considerável mobilização de recursos cognitivos para (re)produção de discursos similarmente situados para exemplificar a possibilidade de uso das funções confluentes com a listagem das autoras.

As interações 6, 7 e 9 foram categorizadas como limitadas por não se calcarem num processo de bilateral, e sim unilateral. Dessa forma, a co-construção do discurso ocorre de maneira limitada, mediada pelo avaliador, o que não é necessariamente negativo ao ambiente da prova, mas certamente afasta-se de co-construções de âmbito social. Assim afirmamos porque, segundo Taylor (2011), a função do examinador é maximizar as oportunidades de produção de um candidato, não necessariamente co-construir suas opiniões. Isto é, cremos ser improvável a presença de mediadores-avaliadores que interfiram negativamente (ex: desafiando uma opinião cujo mediador-avaliador discorda) na produção do candidato.

A interação 7 ocorre por meio de perguntas como a número cinco da parte 2, em que uma justificativa explícita é exigida. No entanto, não há critério na grade que avalie a completude das tarefas, salvo pela parte de recursos lexicais que menciona vocabulário “preciso” e “flexível” em todos os assuntos elencados. No entanto, para a simulação e mobilização de contexto e construto acadêmico sociocognitivamente alinhados, além das funções sociais suscitadas nas tarefas, devemos pensar na natureza do conteúdo esperado (INOUE et al, 2018). No entanto, tal aspecto dos exames orais ainda é ponto de discussão entre especialistas.

A interação 9 é permitida ao candidato em qualquer momento do exame caso ele não entenda o que foi dito pelo examinador, o que configura uma negociação de significados nos moldes de Donato (1994). É um tipo de interação característica de ambientes acadêmicos, mas a duração de tal característica pode ser variável em contexto real, principalmente para assuntos de alta complexidade.

Por fim, tecemos comentários sobre as funções voltadas para a função de gerenciamento da interação nas tarefas 1, 2, e 3. As tarefas de gerenciamento 1 e 2 normalmente são mediadas pelo mediador-avaliador, pois o início e a troca de assuntos depende de sua ação para que ocorram. No entanto, nada impede o candidato, por exemplo, de se desviar dos tópicos sugeridos pelo avaliador ou, ao menos, distorcê-los. Afinal, as relações de coerência e coesão são tão dêiticas quanto à escolha de elementos que podem emergir numa interação. O tópico provavelmente pode ser restabelecido pelo mediador-avaliador, contanto que haja compreensão do que foi dito, o que tornaria o ato de “mudar de assuntos” mais complexo. Contudo, essa imposição não apagaria a tentativa de mudança do evento, o que caracteriza passagem parcial de um modo de interação a outro.

É válido ressaltar que a vontade daquele com mais poder hierárquico prevalece na interação, aspecto típico do contexto de exames de proficiência, sendo esse, normalmente, o mediador-avaliador. O que apontamos é que, por conta de elementos como lapsos, ou diferentes percepções de um mesmo tópico, um assunto pode ser conduzido de maneira diferente da estabelecida no esquema mental do mediador-avaliador e, por isso, ser considerado mudança de assunto.

Quanto à função de gerenciamento 3, apesar de Ducasse e Brown (2011) não terem listado a presença de nenhuma, é possível inferir que um candidato pode indicar o fim de seu turno por meio de um comentário como “É isso.”, que atenderia a ideia de conclusão de argumento; ou, simplesmente, ao responder perguntas como a Q1.3, na Quadro 31, o usuário poderia elencar maior número de vantagens do que desvantagens em seu discurso. A imprevisibilidade do desenvolvimento e gerenciamento do assunto depende das capacidades interacionais do usuário da língua, o que nos fez questionar a categorização de Ducasse e Brown (2011) nesses aspectos.

Assim sendo, uma versão do quadro resumo de Ducasse e Brown (2011) investigado sob nossa ótica interpretativa alteraria o resultado original, que conta com seis das dezesseis

funções elencadas, para um com treze. No entanto, concordamos com as autoras quanto à natureza interacional da co-construção do discurso em ambiente acadêmico, isto é, os aspectos de gerenciamento da interação em “boa-fé”, que são cruciais à manutenção da interação, bem como às potenciais rotinas necessárias para realizar todas as demais funções informacionais e interacionais já listadas ao longo deste trabalho (ex: polidez, distância pessoal, tamanho dos turnos de fala, etc.). Resta, então, para objetivos futuros, uma pesquisa que aponte para a relação existente entre proficiência, propósito e frequência de utilização de funções academicamente situadas, pois em nossa atual não temos como prover certezas além de nossas percepções da literatura específica e possibilidades hipotéticas de desenvolvimento da produção.

Quanto ao alinhamento da grade de avaliação às tarefas e ao propósito do exame, notamos que não há critérios tão explicitamente voltados para a avaliação de funções interacionais às seções de léxico e estrutura, que se chamam, respectivamente, “recursos lexicais” e “alcance gramatical e precisão”. Com base nisso, investigamos, nos Quadros 7 e 34, a seção de “fluência e coerência”, a seção restante.

Em princípio, interpretamos a seção como ambígua, pois, não pudemos deixar de notar a presença do termo “fluência”, que é abrangente. “Fluência”, em nossa concepção, era tida um aspecto prosódico da fala, mais comumente relacionado à fonologia, tal qual Weir (2005) aponta como um dos critérios tomados como forte indicador de proficiência em exames orais, o que o torna relativamente desalinhado à ideia de aspectos funcionais.

Contudo, concebemos que a dificuldade de desenvolver um assunto fluente e coerentemente pode ser decorrente de competências que estão além das fonético-fonológicos como a falta de domínio do assunto, falta de familiaridade com a interação, dentre outras, aspecto o qual é considerado no descritor 9 da seção de “fluência e coerência”. Assim sendo, associamos a seção de “fluência e coerência” da grade à esta seção de aspectos funcionais.

Bereiter e Scardamalia (1987) postulam que as demandas sociocognitivas, isto é, as funções discursivas referentes a um contexto, devem estar contidas nas funções das tarefas que os candidatos devem desempenhar. Logo, o alinhamento da grade de avaliação deve ser investigado por meio de um critério que reflita tal qualidade da produção.

Em nossa investigação, as frases e termos da grade de avaliação que mais alinham-se ao que postula Bereiter e Scardamalia (1987) são: 1) a maneira como o candidato fala e utiliza recursos coesivos; 2) o desenvolvimento de assuntos de forma completa e apropriada; 3) a frequência de repetição ou correção; 4) a presença de turnos longos com ou sem hesitação referente ao assunto e; 5) variabilidade de conectivos e marcadores discursivos.

O critério número 1 encontra-se na mesma situação contextualmente complexa que os critérios investigados em léxico e estrutura. Retomando um pouco dessa discussão em que apontamos que a avaliação qualitativa da produção sob a visada de termos como “precisão”, “completamente”, “variado” e “flexivelmente” depende da definição de um contexto a ser simulado. Um exemplo de tal análise está presente neste capítulo, que é na análise de Ducasse e Brown (2011).

Ducasse e Brown (2011) estabelecem os paradigmas contextuais sob os quais desenvolveram sua análise, isto é, um contexto metodológico, ideológico, avaliativo, linguístico e acadêmico. Para poder entender o que as autoras conceberam nas terminologias de suas análises, como no caso do termo “elaborar”, cabe-nos tentar interpretar o texto original sob a visada das autoras.

Após inferir o “ponto de partida”, isto é, as concepções terminológicas utilizadas, procuramos outras vias de investigação, por meio da utilização de diferentes paradigmas, para explorar as diferentes significações que uma palavra pode representar (VAN DIJK, 2012; BAKHTIN, 2017; 2019; VOLOCHINOV, 2018) e, assim, melhor compreender como a subjetividade e objetividade da análise são contextualmente definidas em um documento de investigação.

Apesar das autoras terem feito uma investigação cujo respaldo é notável, a mudança de ponto de vista nos gerou contribuições e *insights* que apontaram para a inerente necessidade de uma definição clara do que deve ser avaliado. Em Ducasse e Brown (2011), não era somente o caráter das funções acadêmicas que fazia parte da investigação, mas tal como em Inoue et al (2018), as características informacionais, além dos aspectos formais da língua, que compunham a resposta.

Para que se estabeleça maior relação entre validade de contexto, tarefas e grade de avaliação, é necessário, assim como constatado nos aspectos lexicais e estruturais, definir a

forma adequada de se desenvolver um assunto, que definitivamente não podem ser definidos como “flexivelmente”; logo, a relevância interacional (WEIR, 2005) do assunto em relação ao propósito da prova e expectativa de confluência de funções que podem ser concatenadas para que o desenvolvimento de um assunto possa ser considerado relevante deve sempre ser mantida como ponto focal.

Apesar de haver um misto de funções cujo caráter varia entre simples e complexo em cada uma das partes do exame oral do IELTS, os aspectos funcionais e lexicais (LUOMA, 2004; WHITE, 2020) que são pontuadas por meio da grade de avaliação não nos parece conter aspectos claros quanto às diferenças entre as expectativas de produção de candidatos ao exame geral e acadêmico, o que por si só, é um problema. O que é desenvolver um tópico completamente de maneira geral? Como tal critério se aplica ao contexto acadêmico? Claramente, essas perguntas são retóricas. Apesar de estarmos tratando de outro assunto, certas partes da análise se repetem.

Além de haver, mais uma vez, notável sobrecarga cognitiva pairando sobre o julgamento do mediador-avaliador para avaliar candidatos de maneira justa, não há nada que aponte-o em direção a uma análise de conteúdo, ou expectativa de resposta mais apropriada do que a investigação das funções necessárias à cada um dos contextos suscitados (acadêmico e geral) e o significado contextual dos termos listados na grade: completamente, variado, flexível, etc.

Pensando nos aspectos coesivos e desenvolvimento do assunto, o domínio da seção “fluência e coerência”, alvo de nossa investigação dos aspectos funcionais do exame, parece ligar-se mais aos recursos lexicais e estruturais utilizados na produção do que o domínio das rotinas interacionais por si só. Observe a lista retirada da referida seção da grade de avaliação (SEEDHOUSE; EGGBERT, 2004): recursos coesivos; o desenvolvimento de assuntos de forma completa e apropriada; frequência de repetição ou correção; presença de turnos longos com ou sem hesitação referente ao assunto; variabilidade de conectivos e marcadores discursivos.

Salvo pelos “recursos coesivos”, “conectivos” e “marcadores discursivos”, que possuem, certamente, caráter lexical com funções prevalentes e alinhadas à avaliação dos aspectos funcionais de um determinado desempenho, os demais itens, parecem fazer maior

interface com as outras seções da grade de avaliação e seus respectivos critérios de avaliação do exame.

Não questionamos a organização dos critérios somente pela natureza semântica dos termos, mas também pela falta de especificidade do que se entende como “coerência” na grade de avaliação, dado que o fio temático da prova é estabelecido somente nas partes 2 e 3. Como avaliar a coerência da parte 1, por exemplo? Ainda que a distribuição dos pesos da avaliação seja holística, favorece-se a avaliação temática das partes que sucedem a primeira, mais do que a primeira em si.

5.16 Considerações finais da análise

Para encerrar o capítulo, o que vemos como fundamental à avaliação e seus aspectos inerentemente voltados à validade de contexto é a forma com que as tarefas mobilizam os construtos e o caráter funcional, lexical e estrutural dos construtos-alvo ao tentar conectá-los a aspectos de alta relevância ao propósito do exame. Ademais, transparência e coerência nos critérios da grade de avaliação são fundamentais para que esses possam adequar-se às competências necessárias para a mobilização de recursos cognitivos que auxiliem na produção de interações cuja natureza esteja alinhada ao contexto.

Considerações Finais

Os resultados da pesquisa, além de proporcionarem respostas às perguntas dessa investigação, levaram a considerações além da base bibliográfica a que utilizamos. A compreensão sobre uma abordagem que envolva a sociocognição é complexa e agora entendemos com mais exatidão as linhas tênues que permeiam o contexto dos exames de proficiência.

Conseguimos constatar o desalinhamento do exame oral do IELTS em relação à modalidade acadêmica por meio do parcial desvelamento do que cremos serem aspectos de um construto de fala acadêmico com base na investigação da validade de contexto. Acreditamos que, por calcar-se num modelo generalista de proficiência, em vez de um modelo cujas funções acadêmicas sejam mais avidamente mobilizadas, o exame oral do IELTS não prevê funções sociocognitivas do contexto acadêmico como poderia.

Apesar da análise dos critério de Ducasse e Brown (2011) ter nos levado à conclusão parcial de que o exame oral do IELTS tem potencial para mobilizar treze das dezesseis características fundamentais de um contexto acadêmico, a operacionalização se torna inerentemente dependente da estruturação das tarefas e da proficiência do candidato, e não só, necessariamente, da maneira com que o exame avalia (apesar desse critério também ter se mostrado insuficientemente posta em termos de completude da resposta) os construtos referentes ao contexto acadêmico.

Além disso, Inoue et al (2018) elencam uma característica fundamental ao construto acadêmico, a qual foi comentada quanto às pesquisas de Ducasse e Brown (2011), que é referente à qualidade da elaboração das respostas: a análise de conteúdo da resposta. A relação de complexidade estabelecida entre temas concretos e abstratos pode ser suficiente para temas gerais, mas, se tratando do formato retórico acadêmico, isto é, as exigências discursivas e intelectuais que emergem em tal contexto, não se pode afirmar que a adequação de funções, temas e tarefas de um exame de proficiência geral seja suficiente para mobilizar um construto de caráter acadêmico.

Essas descobertas investigativas nos levam a entender o porquê um mesmo modelo de prova é utilizado para duas modalidades com propósitos diferentes e amplia nossa

compreensão de como a validade de contexto pode ser estabelecida em um exame, mas não pudemos deixar de notar os desalinhamentos presentes na grade de avaliação do exame, assim como em sua estrutura como um todo. Ao passo em que isso é algo notável, como já citado na introdução, é também, normal; afinal, não existe um exame que seja completamente adequado aos seus propósitos, e o valor desta dissertação encontra-se, principalmente, nas contribuições feitas à estrutura do IELTS.

Nossas considerações finais, sob a lente sociocognitiva e ante os resultados obtidos nesta dissertação, é que a validade de contexto das tarefas do exame oral do IELTS tem alinhamento parcial com a grade avaliativa e seus propósitos. Parcial, pois, como parte do propósito é geral, há adequação de contexto. No entanto, como o exame oral do IELTS também considera a modalidade acadêmica por meio do mesmo modelo de exame, notamos que, não só nas funções avaliadas da grade de avaliação, quanto na estrutura, léxico, função, interação, e tarefas, são necessários ajustes para melhor mobilizar as demandas sociocognitivas do contexto acadêmico.

Sobre nossos objetivos gerais, cremos que todos foram alcançados, mas não cremos que essa investigação se encerre neste trabalho. A investigação das tarefas e enunciados demonstrou alinhamento com propósitos gerais, além de ter revelado que a interação dominada pelo avaliador tende a reduzir funções interacionais relevantes à qualidade da informação exigida no meio acadêmico. A análise da grade avaliativa investigou grandes incongruências terminológicas, uma vez que o que define-se como “flexível” em ambiente acadêmico e geral, por exemplo, define-se por instâncias pragmáticas diferentes, impossibilitando uma análise de desempenho com potenciais critérios subjetivos e cognitivamente exigentes para o mediador-avaliador.

Outrossim, as relações entre as tarefas, enunciados, comandos e a grade de avaliação não estão adaptadas aos diferentes contextos acadêmicos e gerais. Nossa investigação revelou que as instruções, enunciados e comandos das tarefas estão voltados aos níveis mais baixos de proficiência, o que por consequência, pode inibir a performance de candidatos de nível mais avançado de maneira autêntica, uma vez que, nos exemplos analisados, a interação teve que ser provocada. Dessa forma, ainda que consideremos os resultados parcialmente alinhados aos propósitos gerais da prova, não os consideramos apropriados para as funções acadêmicas. Consideramos que as contribuições mais marcantes de candidatos acadêmicos viriam por

meio da contribuição interacional entre cada um deles, por meio de atividades que os permitissem utilizar funções mais complexas da linguagem internacionalmente, assim como reparação de comunicação, elaboração das ideias dos colaboradores do exame, dentre outros fatores.

Em relação às perguntas elencadas por Weir (2005), resumidas no final do capítulo 3, eis um apanhado de respostas contendo pontos nevrálgicos de nossa análise.

Perguntas de Weir (2005)	Respostas de nossa investigação
<i>[A](s) [instruções, enunciados e] comando(s) do teste são preciso(s) e acessível(is)?</i>	As instruções, enunciado e comandos foram desenvolvidos pensando em múltiplos níveis e mesclam diferentes funções interacionais. Logo, vemo-os como possivelmente alinhados a múltiplos níveis, mas não necessariamente a todos os propósitos da prova.
<i>O objetivo do teste é inequivocamente claro para o candidato? É um propósito apropriado?</i>	Sim, para propósitos gerais, e não para propósitos acadêmicos. A relação entre contexto e uso da língua não se manifesta de maneira apropriada para o segundo caso.
<i>Existe alguma evidência de que o formato de resposta do teste pode afetar o desempenho de seu candidato?</i>	Sim, como citado em Seedhouse e Morales (2017), a adequação ao formato majoritariamente monológico do teste gera certa estranheza e deixa de mobilizar interações que possam favorecer a produção do candidato.
<i>Os critérios a serem usados para pontuar as tarefas da prova são explícitos para os candidatos e os examinadores?</i>	Os critérios são acessíveis, mas não necessariamente explícitos. Ajustes quanto à clareza dos critérios são necessários.
<i>De que maneira um item que testa um aspecto X é mais ou menos importante do que um item que testa um aspecto Y? Há alguma ponderação para os diferentes componentes de um teste que tenha sido devidamente justificada?</i>	O exame avalia todas as suas partes igualmente. A justificativa para suas diferenças revela em termos de dificuldades é revelada na seção referente à ordem das tarefas, que foram ajustadas para acomodar o candidato ao formato da prova (GALACZI; FFRENCH, 2011)
<i>Os itens e as tarefas de um teste estão em uma ordem justificável?</i>	A ordem é justificável com base no fato de que o exame avalia múltiplos níveis, mas não necessariamente alinhada a múltiplos propósitos.
<i>O tempo para cada parte do teste, por exemplo, preparação e conclusão, é apropriado?</i>	A divisão do tempo é relativamente igualitária, com a parte 1 variando de 3 até 5 minutos, a 2 variando de 2 a 4 minutos e a 3 variando de 3 a 5 minutos; no entanto, não conseguimos fazer observações sobre a organização e influência da divisão do tempo, exceto que candidatos com propósitos acadêmicos teriam mais chance de mobilizar recursos cognitivos contextualmente alinhados com mais facilidade na parte 3.

<i>O modo de discurso é, afinal, apropriado para as habilidades ou estratégias que se deseja avaliar?</i>	Sim, para o propósito geral, e potencialmente adequado para candidatos acadêmicos de alto nível de proficiência, mas não completamente apropriado de acordo com as pesquisas apresentadas no início do capítulo de análise e a última seção dos aspectos funcionais do exame.
<i>O canal é apropriado para os requisitos de situação-alvo dos alunos que estão sendo testados?</i>	Parcialmente apropriado para interações acadêmicas, podendo beneficiar-se da inserção de interações face-a-face entre candidatos.
<i>A extensão do texto é apropriada para os requisitos da situação-alvo dos alunos que estão sendo testados?</i>	De acordo com nossa investigação, considerando texto como os diferentes insumos auditivos fornecidos pelo mediador-avaliador e pelo cartão da parte 2, concluímos que o léxico e estrutura presentes nas instruções, enunciados e comandos das tarefas encontram-se de acordo com níveis que variam do A2 ao B1, servindo, teoricamente, para quaisquer níveis de proficiência superiores a esses. A extensão certamente não atrapalha candidatos de baixa proficiência, mas nosso referencial revela que candidatos de alta proficiência podem ter seu desempenho afetado.
<i>O tipo de informação é apropriado para os requisitos da situação-alvo dos alunos que estão sendo testados?</i>	A informação é de caráter geral, e varia entre tarefas que exigem informações abstratas e/ou concretas, mas não podemos afirmar se isso é necessariamente prejudicial ao contexto acadêmico. As pesquisas que utilizamos para investigar as funções cognitivas tidas como acadêmicas o faziam em ambientes pedagógicos, analisando, tão-somente, o caráter interacional da informação, e não sua natureza concreta ou abstrata. Acreditamos que exames orais cuja função seja acadêmica deve tratar de processos organizacionais mais complexos para maximizar a maior quantidade de amostras possíveis.
<i>O assunto está de acordo com a experiência e o nível provável da faixa etária dos alunos?</i>	Os assuntos que supomos serem possíveis parecem apropriados para múltiplas audiências (LUOMA, 2004; WHITE, 2020). As diferentes funções discursivas distribuídas em cada uma das perguntas das tarefas torna a adequação para múltiplos níveis possível, mas, para alunos mais avançados, pode limitar sua produção.
<i>O assunto é adequado para todos os alunos? O assunto não deve ser tendencioso de nenhuma forma (em termos de urbano/rural, menino/menina, cultural, etc.).</i>	A parte 3 do exame é a parte que mais se adequa ao que Weir (2005) postula sobre a adequação da familiaridade de uma questão de um exame de proficiência. Os assuntos ali contidos tendem a ser familiares, mas não tanto a ponto de permitirem a mobilização de esquemas mentais complexos. O maior diferencial dessa seção do exame encontra-se nas tarefas que suscitam habilidades cognitivas do candidato de afastar-se da realidade imediata imposta pelas condições materiais para formular hipóteses, ou especulações de cunho abstrato.
<i>O estilo de diálogo/passagem é familiar aos alunos? O tipo de texto deve ser semelhante ao do livro didático, por exemplo, descrição, narrativa, argumento.</i>	O estilo de diálogo pode ser familiar aos candidatos, mas está mais próximo de uma entrevista, ou interrogatório, do que de interações acadêmicas. O turno longo de fala possui temática que talvez não se relacione com ambientes acadêmicos comuns (tópicos cotidianos), além de exigir poucas habilidades linguísticas abstratas. Em relação à competência geral de fala, há certa familiaridade com vários ambientes pragmaticamente funcionais calcados em gêneros discursivos socialmente fundamentados (ex: relato experiencial) tal como Van Dijk (2012).

O conteúdo do assunto é apropriado para os requisitos da situação-alvo dos alunos que estão sendo testados?

Para os que prestam para a modalidade geral, sim. Para os que prestam para a modalidade acadêmica, não inteiramente. Todavia, não há como afirmar que os assuntos são pertinentes ao ambiente acadêmico com tanto afinco com que podemos afirmar que eles estão alinhados à contextos gerais de proficiência do uso do inglês.

Os itens lexicais no teste estão no texto de insumo e são exigidos na produção de maneira adequada ao nível dos candidatos?

No insumo, os itens lexicais apresentam relativa adequação aos propósitos, ainda que haja espaço para melhorias que tragam mais funções acadêmicas à tona. Na grade de avaliação, no entanto, muitos aspectos permanecem vagos e carecem de definições contextualmente estabelecidas que, nem o exame e nem a grade provêm. Isso acaba por tornar o critério subjetivo e delega muita responsabilidade ao mediador-avaliador. Ademais, muito não é explicado sobre o que vem a ser o uso preciso de situações do léxico, uma vez que isso é julgado de acordo com o assunto apresentado no exame, o que reforça a prevalência do julgamento do avaliador mais do que dos instrumentos investigados.

Os itens gramaticais no texto estão no texto de insumo e os exigidos como produção (por parte da grade de avaliação) são apropriados para o nível dos candidatos?

Por se tratar de um exame que afere múltiplos níveis de proficiência, os itens gramaticais não variaram, e cremos que não variam, muito além das estruturas verbais mais básicas. Isso torna o aspecto estrutural associado em termos de insumo. Contudo, novamente, na grade de avaliação, muito há de ser explicado quanto aos critérios apresentados, além de haver necessidade de uma melhor definição de uso “preciso”. Seria o uso preciso contextualmente definido, ou prescritivamente definido ao contexto? Sendo assim, o que as organiza em termos gerais e acadêmicos? Erudição? Curso acadêmico? Cabe ao exame especificar algumas diretrizes para assegurar maior validade de contexto.

As funções elencadas no teste estão claras no texto de insumo e são exigidas como produção de maneira apropriada para o nível do candidato?

As funções elencadas nas tarefas são cabíveis para ambas as modalidades, acadêmica e geral, mas poderiam ser melhor organizadas quanto à exigência de gerenciamento internacional. Tal como Goffman (1967) aponta, cada contexto apresenta sua própria estrutura ritualística, com fundamento social (VAN DIJK, 2012) e que talvez possa ser melhor generalizada (LUOMA, 2004) para acomodar ambas as modalidades.

Contudo, no formato atual, pensando na relação entre as tarefas e os critérios da grade de avaliação, não podemos dizer que há como garantir um exame cuja validade de contexto se sustente.

Em resumo, os instrumentos não estão alinhados aos seus propósitos, o que afeta sua validade de contexto. Dessa forma, sua validade de contexto nos parece parcial para o contexto acadêmico, pois tanto as tarefas quanto a grade de avaliação parecem voltar-se para amostras em que o desempenho geral é priorizado. Como as tarefas e a grade de avaliação não refletem dois construtos diferentes, uma avaliação justa do desempenho dos candidatos se torna complexa, mas ao mesmo tempo, favorece a praticidade de aplicação da prova.

À luz de nossos paradigmas sociocognitivos, então, concluímos múltiplos desalinhamentos, ainda que não totais, à modalidade acadêmica. Em nossa visão, a grade de avaliação precisa de critérios e seções voltadas ao desempenho acadêmico, ainda que em contexto acadêmico geral, para que possa ser mais adequadamente alinhada ao seu contexto de maneira sociocognitivamente válida. Uma sugestão válida, que pode ser desenvolvida em outras pesquisas, é uma proposta de junção, em vez de categorização de critérios inerentes à habilidade fala acadêmica por função e nível de complexidade contextual nos descritores (ex: função de fala X, Y e Z junto do uso de conjunções de função X, Y e Z). Basicamente, se tivéssemos que organizar uma grade de avaliação com os dados que analisamos hoje, partiríamos de três critérios: funções interacionais - que avaliaria a mobilização de funções discursivas inerentes ao contexto acadêmico (BROWN; YULE, 1983; BROWN, ANDERSON, SHILLCOCK; YULE, 1984; BYGATE, 1987; O’SULLIVAN, 2002; DUCASSE; BROWN, 2011; INOUE ET AL, 2018); qualidade da informação (DUCASSE; BROWN, 2011; VAN DIJK, 2012) que visaria traçar diferenças entre critérios vagos, tais como "flexivelmente" e "completamente", com base na elaboração retórica do contexto pertinente ao candidato de forma geral, ou específica, a depender das especificações do exame; por fim, a capacidade abstração (BYGATE, 1987; COUNCIL OF EUROPE, 2011) seria o último critério que poderia ser relacionado com os outros dois anteriores por meio da avaliação das relações traçadas entre os comentários produzidos pelo candidato e sua realidade imediata. As tarefas, tal como Ducasse e Brown (2011) e Inoue et al (2018) apontam, dependem de pesquisas que gerem dados para um melhor entendimento do que pode ser necessário aos contextos de candidatos acadêmicos. Cabe ressaltar que o conteúdo das tarefas é de suma importância, mas que tal assunto merece consideração mais focal em investigações de testes de conteúdos específicos, como o GMAT.

Carregamos como aprendizado que o referencial teórico para pesquisas de cunho sociocognitivo é extremamente denso. A necessidade de uma costura teórica entre linhas acadêmicas diferentes, mas relacionadas, foi uma das maiores demandas de leitura que tive na vida, mas que me proporcionou diferentes visões de um mesmo ponto de análise para investigação. Outro ponto que merece atenção é o rigor metodológico que estudos de avaliação tendem a receber. Cremos que, para pesquisas futuras no mesmo tema, em termos metodológicos, uma análise de corpus seria capaz de beneficiar a investigação, principalmente

porque nos auxiliaria a estabelecer um construto acadêmico geral por meio de nossas próprias investigações.

Aos demais pesquisadores que possam vir a se interessar pelo campo com base em nossas contribuições, olhar para os fundamentos de cada uma das bases de nossos referencial é fundamental. É necessário entender, ao menos um pouco, de pedagogia, linguística, neurocognição, metodologias de ensino e diferentes escolas linguísticas para uma análise que parta de diferentes pontos de vista.

Referências

- ALDERSON, J.C. *Assessing Reading*. Cambridge: Cambridge University Press, 2000.
- _____; CLAPHAM, C.; WALL, D. *Language Test Construction and Evaluation*. Cambridge: Cambridge University Press, 1995.
- AMARAL, D. *A perspectiva dos examinadores sobre o uso da grade de avaliação do IELTS*. Dissertação (mestrado) - Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Linguística Aplicada. Porto Alegre, 2011.
- BACHMAN, L. F. *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press, 1990.
- _____. PALMER, A. *Language Testing in Practice: Designing and Developing Useful Language Tests*. USA: OUP, 1996.
- _____. Modern language testing at the turn of the century: Assuring that what we test counts, *Language Testing*, n. 17, p. 1–42, 2000.
- Bakhtin, M.. *Teoria do romance I: A estilística*. Tradução: Paulo Bezerra. 2ª. ed. São Paulo: editora 34, 2017
- _____. *Questões de estilística no ensino da língua*. Tradução: Paulo Bezerra. 2ª. ed. São Paulo: editora 34, 2019
- BALL, F. Developing wordlists for BEC. In: *Research Notes 8*. Cambridge ESOL, pp. 10-13, 2002.
- BARDIN, L. *Análise de conteúdo*. Lisboa: Ed. 70, 1979.
- BEREITER, C.; SCARDAMALIA, M. *The Psychology of Written Composition*. Hillside, New Jersey: Lawrence Erlbaum Associates, 1987.
- BROOKS, L.; SWAIN, M. *Contextualizing performances: Comparing performances during TOEFL iBT and real-life academic speaking activities*. *Language Assessment Quarterly* 11, 2014
- BROWN, A. *An investigation of raters' orientation in awarding scores in the IELTS interview* in TULLOH, R. (Org.) *IELTS Research Reports*. Volume 3 IDP: Camberra , 2000

BROWN, H. D. *Teaching by principles: an interactive approach to language pedagogy*. New York: Longman, 2007.

BROWN, G.; YULE, G. *Discourse Analysis*. Cambridge: Cambridge University Press, 1983.

_____. ANDERSON, A.; SHILLCOCK, R.; YULE, G. *Teaching Talk: Strategies for Production and Assessment*. Cambridge: Cambridge University Press, 1984

BUCK, G. *Assessing Listening*. Cambridge: Cambridge University Press, 2001.

BYGATE, M. *Speaking*. Oxford: Oxford University Press, 1987.

_____. *Effect os task repetition on the structure and control of oral language*, in BYGATE, M.; SKEHAN, P.; SWAIN, M. *Researching Pedagogic Tasks: Second Language Learning, Teaching and Assessment*. Harlow: Longman, 2001.

CAPPELLA, J. N. (1996). *Dynamic coordination of vocal and kinesic behavior in dyadic interaction: Methods, problems, and interpersonal outcomes*. In WATT, J; VANLEAR, C. (ORG.). *Dynamic patterns in communication processes*. Thousand Oaks, CA: Sage.

CARROL, J. B. *Fundamental Considerations in testing for English language proficiency of foreign students* in *Testing the English Proficiency of Foreign Students*. Washington: Center for Applied Linguistics, 1961

CEFR. *Common European Framework of Reference*. Council of Europe, 2001.

CEFR. *Common European Framework of Reference*. Council of Europe, 2018.

CEGALLA, D. P. *Novíssima gramática da língua portuguesa*. 48. Ed. São Paulo: IBEP, 2009.

CHAPELLE, C. *Construct definition and validity inquiry in SLA research*. In BACHMAN, L.F.; COHEN, A.D. (Org.), *Interfaces between Second Language Acquisition and Language Testing Research*. New York: Cambridge University Press, 1988.

CHESIRE, J. *Variation in the use of ain't in an urban British English dialect*. Cambridge University Press: 2008.

COLE, M.; COLE, S. R. *The development of children*. 3ª ed. New York: W. H. Freeman, 1996.

COOPER, R. L. *An elaborated language testing model* in *Language Teaching* Special Issue 7, 1968

CRESSWELL, J. W. *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. New York, Boston: Pearson, 2002.

DENZIN, N. K.; LINCOLN, Y. S. *Introduction: The discipline and practice of qualitative research*. In: _____; _____ (Eds.). *The Sage Handbook of qualitative research*. 4. ed. Thousand Oaks: Sage, 2005.

DOISE, W.; MUGNY, G. *Psychologie sociale et développement cognitif*. Paris: Armand Colin, 1997.

DONATO, P. *Collective Scaffolding*. In: LANTOLF, J. P. (org.). *Vygotskyan Approaches To Second Language Research*. New York: Ablex Pub Corp, 1994.

DOUGLAS, D. *Assessing Languages for Specific Purposes*. Cambridge: Cambridge University Press, 2000.

DUCASSE, A. M.; BROWN, A. *The role of interactive communication in IELTS speaking and its relationship to candidates' preparedness for study or training contexts*. Editor IDP: IELTS Australia and British Council, 2011

ELLIS, R. *Constructions, chunking and connectionism* In: DOUGHTY, C.; LONG, M. (Org.) *The Handbook of Second Language Acquisition*, Oxford: Blackwell, 2003.

FERNANDES, I. S. (Org). *Manual de Orientações para a Elaboração de Itens: Orientações técnico-pedagógicas para o elaborador*. Belo Horizonte: Faculdades Promove; Infórium; Kennedy. 2015. Disponível em: <
http://www.kennedy.br/arquivos_up/documentos/2175911dacec54774d363e3311f04a66.pdf >
 Acesso em: 16 de Julho 2021.

FIELD, J. *Cognitive validity*. In: TAYLOR, L. (Ed.) *Studies in Language Testing* 30. Cambridge: UCLES/Cambridge University Press, 2011.

FIELDING, N.; SCHREIER, M. *Introduction: On the Compatibility between Qualitative and Quantitative Research Methods*. In: *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research* 2001 (revista on-line) Disponível em: <
<http://qualitative-research.net/fqs/fqs-eng.htm> > acesso em 20 de junho de 2021.

- FONSECA, J. J. S. *Metodologia da pesquisa científica*. Fortaleza: UEC, 2002.
- FREITAS, M. T. A. *Vygotsky e Bakhtin - Psicologia e Educação: Um Intertexto*. 4a. ed. São Paulo: Ática, 2003.
- FULCHER, G.; DAVIDSON, F. *Language Testing and Assessment*. London, New York: Routledge, 2007.
- GALACZI, E.; FFRENCH, A. Context validity. In: TAYLOR, L. (Ed.) *Studies in Language Testing 30*. Cambridge: UCLES/Cambridge University Press, 2011.
- GERHARDT, T. E.; SILVEIRA, D. T. (org). *Métodos de Pesquisa*. Coordenado pela Universidade Aberta do Brasil – UAB/UFRGS e SEAD/UFRGS. Porto Alegre: Editora da UFRGS, 2009.
- GIBBS, G. *Análise de dados qualitativos*. Porto Alegre: Artmed, 2009.
- GOFFMAN, E. *Interaction Ritual: essays on face-to-face behavior*. Penguin University Books, 1967.
- _____. *Forms of talk*. England: Basil Blackwell Publisher, 1981.
- _____. *Frame analysis: an essay on the organization of experience*. Boston: Northeastern University Press., 1986.
- GOVERNO DO ESTADO DE MINAS GERAIS - SECRETARIA DO ESTADO DE EDUCAÇÃO. *Guia de Elaboração e Revisão de Itens de Múltipla Escolha*. Online: Disponível em <
http://www.adventista.edu.br/_imagens/area_academica/files/guia-de-elaboracao-de-itens-120804112623-phpapp01%283%29.pdf> Acesso em: 16 de Julho de 2021.
- GUMPERZ, J. J. *Language in Social Groups. Essays Selected and Introduced by Anwar S. Dil. Stanford*. California: Stanford University Press, 1971.
- _____; HYMES, D. *Directions in Sociolinguistics: The Ethnography of Communication*. New York: Holt, 1972.
- GRAY, D. E. *Pesquisa no mundo real*. 2.ed. Porto Alegre: Penso, 2012.
- _____. *Discourse Strategies*. New York: Cambridge University Press, 1982.

HALLIDAY, M.A.K, *Explorations In The Functions Of Language*. New York: Elsevier, North-Holland, 1973.

_____. *Anti-Languages in The Anthropology of Language*, vol. 78. 1976.

HAWKINS, J. A.; BUTTERFLY, P. *Criterion Features in Learner Corpora: theory and illustration*. Cambridge: CUP, 2009

HAZENBERG, S.; HULSTJIN, J. H. *Defining a minimal receptive second language vocabulary for non-native university students: an empirical investigation*. *Applied Linguistics*: 1996.

HINDMARSH, R. *Cambridge English Lexicon*. Cambridge: CUP, 1980

HYMES, D. H. On communicative competence. In: PRIDE, J.B.; HOLMES, J. *Sociolinguistics: selected readings*. Harmondsworth: Penguin, 1972.

_____. *Foundations of Sociolinguistics: An Ethnographic Approach*. Philadelphia: University of Pennsylvania Press. 1974.

IELTS, 2021. *Test Structure*. Disponível em: < <http://www.ielts.org/>>. Acesso em: 20 de maio de 2021.

INGRAM, D.; BAYLISS, A. *IELTS as a predictor of academic language performance*. IELTS Australia and British Council: Camberra, 2007.

INOUE, C.; NAKATSUHARA, F.; LAM, D. M. K.; TAYLOR, L. *Academic speaking: does the construct exist, and if so, how do we test it?* in CRELLA Spring Research Seminar: Luton, 2018.

JAPIASSÚ, H.; MARCONDES, D. *Dicionário Básico de Filosofia*. 5.ed. Rio de Janeiro: Zahar, 2008.

JOHNSON, M. *A Philosophy of Second Language Acquisition*. London: Yale University Press, 2003.

KETTLE, M.; MAY, L. The ascendancy of oracy in university courses: Implications for teachers and second language users. In: GITSAKI, C.; BALDAUF Jr, R. B. (Org.) *Future Directions in Applied Linguistics: Local and Global Perspectives*. United Kingdom: Cambridge Scholars Publishing, 2012.

KHALIFA, H. (org.) *Student Achievement Test Development Manual (SATD)*. Egypt: USAID, 2003

KIM, U. Factors influencing academic achievement in relational cultures: The role of self, relational and collective efficacy In PAJARES, F.; T. URDAN (Org.) *Self-efficacy beliefs of adolescents*. Information Age Publishing, 2006.

KOCH, I. *Introdução à Linguística Textual*. São Paulo: Martins Fontes, 2004.

KOZULIN, A.; GINDIS, B. *Vygotsky's Educational Theory in Cultural Context*. London: Cambridge University Press, 2003.

KROLL, B. *What does time buy? ESL student performance on home versus class compositions*. In KROLL, B. (ORG.) *Second Language Writing: Research Insights for the Classroom*. Cambridge: Cambridge University Press, 1990.

LADO, R. *Language testing: The construction and use of foreign language tests: A teacher's book*. Bristol: Longmans, Green and Company, 1961.

LANTOLF, J. P. (org.). *Vygotskian Approaches To Second Language Research*. New York: Ablex Pub Corp, 1994.

_____. *Sociocultural theory and second language learning*. London: Oxford University Press, 2000.

LANTOLF, J. P., Thorne, S. L.; POEHNER, M. *Sociocultural Theory and Second Language Development*. PATTEN, B. VAN; WILLIAMS, J. (orgs.). *Theories in Second Language Acquisition*. New York: Routledge, 2015.

_____.; PAVLENKO, A. *Social Cultural Theory and Language Acquisition*. In: *Annual Review of Applied Linguistics*. Cambridge: Cambridge University Press, 1995.

_____.; POEHNER, M. *Sociocultural Theory and the Pedagogical Imperative in L2 Education: Vygotskian Praxis and the Research/Practice Divide*. New York: Routledge, 2014.

LAUFER, B. *What's in a word that makes it hard or easy? Intralexical factors affecting the difficulty of vocabulary acquisition* in MCCARTHY, M.; SCHMITT, N. (Org.) *Vocabulary Description Acquisition and Pedagogy*. Cambridge: CUP, 1997.

- LAZARATON, A.; FRANTZ, R. *An analysis of the relationship between task features and candidate output for revised FCE speaking examination*. UCLES, 1997
- LEONT'EV, A. N. The problem of activity in Psychology. In: WERTSCH, J. V. *The Concept of Activity in Soviet Psychology*. Armonk, New York: M.E Sharpe, 1981.
- LEVINSON, S. *Space in language and cognition: Explorations for cognitive density*. Max Planck Institute for Psycholinguistics: Research Gate, 2003
- LUOMA, S. *Assessing Speaking*. Cambridge: Cambridge University Press, 2004.
- MANHATTAN, 2021 *History of the IELTS*. Disponível em <<https://www.manhattanreview.com/ielts-history/>> Acessado em 20 de maio de 2021.
- MARCONI, M. A.; LAKATOS, E. M. *Fundamentos da Metodologia Científica*. 5ª ed. São Paulo, Campos Elísios: ATLAS, 2003.
- MATTHIESSEN, C. M. I. M.; HALLIDAY, M. A. K. *Systemic functional grammar: a first step into theory*. Higher Education Press. 2009.
- MESSICK, S. *Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning*. American Psychologist: 1995.
- MEURER, J. L. *Integrando estudo de gêneros textuais ao contexto da cultura in Gêneros Textuais - Reflexões e ensino*. 2ª ed. revista e ampliada, Rio de Janeiro: Lucerna, 2006.
- MINAYO, M. C. S. *O desafio do conhecimento. Pesquisa qualitativa em saúde*. São Paulo: Hucitec, 2007.
- MORATO, E. M. *Das funções e do funcionamento da linguagem: um estudo das reflexões de L; S; Vygotsky sobre a função reguladora da linguagem e algumas implicações linguístico-cognitivas para a neurolinguística*; 1991, Dissertação (Mestrado em Linguística) – Instituto de Linguística, Universidade Estadual de Campinas; Campinas.
- MORATO, E. *(In)determinação e subjetividade na linguagem de afásicos: a inclinação anti-referencialista dos processos enunciativos*. Cadernos de Estudos Linguísticos, Campinas, v.41, p.55-74, 12 set. 2011. Disponível em: <<https://periodicos.sbu.unicamp.br/ojs/index.php/cel/article/view/8637001>>. Acesso em: 14

jun. 2021

MILLER, R. *Vygotsky in Perspective*. New York: Cambridge University Press, 2011 .

O'SULLIVAN, B.; WEIR, C. J.; SAVILLE, N. *Using observation checklists to validate speaking-tests tasks*. Language Testing: 19, 2002.

POEHNER, M. DAVIN, K. J., LANTOLF, J. P. Dynamic Assessment. In: SHOHAMY, E. OR, I. G. MAY, S. (orgs.) *Language Testing and Assessment*. 3rd edition. Gewerbestrasse: Springer, 2017

PICCARDO, E. NORTH, B. *The Action Oriented Approach*. A Dynamic Vision of Language Education. Multilingual Matters, Bristol, Blue Ridge Summit. Canadá, 2019.

QUEVEDO-CAMARGO, G. Proficiência linguística e avaliação da produção oral em língua inglesa na educação básica. In: EL-KADRI, M. S.; PASSONI, T. P.; GAMERO, R. (Org.). *Tendências contemporâneas para o ensino de língua inglesa: propostas didáticas para a educação básica*. v. 7. Campinas: Pontes, 2014

REED, S. K. *A structure-mapping model for word problems*. in Journal of Experimental Psychology: Learning, Memory and Cognition. 1987

RICHARDS, J. C; RODGERS, T. S., *Approaches and Methods in Language Teaching*. 2ª edição. Cambridge: Cambridge University Press, 2001.

ROACH, Peter. *English phonetics and phonology: A practical course*. 4. ed. Cambridge University Press: Cambridge, 2009

UNDERHILL, A. *Sound Foundations*. 1ª ed. UK: Macmillan, 1994.

RODRIGUES, A. *A Partitura Invisível: Para uma Abordagem Interactiva da Linguagem*. 2ª ed. Colibris: Lisboa, 2005

ROGOFF, B. *A natureza cultural do desenvolvimento humano*. Porto Alegre : Artmed, 2005.

RAMOS, A. A. L. *Um caminho estrangeiro na compreensão do gênero: estratégias cognitivas em produção textual do Celpe-Bras*. Tese (Doutorado) - Universidade de Brasília. Instituto de Letras, Departamento de Linguística, Português e Línguas Clássicas, Programa de Pós-Graduação em Linguística. Brasília, 2007.

_____ ; VIEIRA, J. A.; SILVA, F. C. O. Teorias Cognitivas e Abordagens Metodológicas no

Ensino de Segunda Língua (L2) e de Língua Estrangeira (LE). In: VIEIRA, J. A.; SILVA, F. C. O. da. (Orgs.). *O que a distância revela - Reflexões de professores e estudantes do Curso de Letras - EaD/UnB*. Brasília: Gráfica e Editora Movimento Ltda., 2014.

ROSENFELD, M.; LEUNG, S.; OLTMAN, P. K.. *The Reading, Writing, Speaking, and Listening Tasks Important for Academic Success at the Undergraduate and Graduate Levels*. Monograph Series. Princeton, New Jersey: Educational Testing Service, 2001.

ROST, M. *Listening in Language Learning*. London: Longman: 1990

SALVADOR, A. D. *Métodos e técnicas de pesquisa bibliográfica*. Porto Alegre: Sulina, 1986.

SEARLE, P. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: CUP, 1969.

SCHMITT, N. *Vocabulary in Language Teaching*. Cambridge: Cambridge University Press, 2000.

SCRIVENER, J. *Learning Teaching*. 3rd edition. London: Macmillan, 2011.

SEEDHOUSE, P.; HARRIS, A.; NAEB, R.; UNSTUNEL, E. *The relationship between speaking feature in IELTS Research Reports Online Series*. vol. 2, UK: New Castle, 2014.

_____.; EGGBERT, M. The interactional organisation of the IELTS Speaking Test. In: *IELTS Research Reports Online Series*. vol.6. UK: Bridgewater House, 2006.

_____.; MORALES, S. *Candidates questioning examiners in the IELTS Speaking Test: An intervention study*. in *IELTS Research Reports Online Series*. vol 5. Australia: IPD, 2017

SHAW, S. D.; WEIR, C. J. *Examining Writing: Research and Practice in Assessing Second Language Writing*, Studies in Language Testing 26. Cambridge: UCLES/Cambridge University Press, 2007.

SKEHAN, P. Tasks and language performance assessment. In: BYGATE, M.; SKEHAN, P.; SWAIN, M.(Org.) *Researching Pedagogic Tasks*. London: LUP, 1997

_____. *A Cognitive Approach to Language Learning*. Oxford: OUP, 1998

_____. FOSTER, P. The influence of planning and post-talk activities pm accuracy and complexity in task based learning. In: *Language Teaching Research* 1, 1997.

SPOLSKY, B. 1995: *Measured words: the development of objective language testing*. Oxford: Oxford University Press.

- SPOLSKY, B. Language Assessment in Historical and Future Perspective. In: *Language Testing and Assessment*. Volume 7, 2008.
- STAKE, R. E. *Pesquisa qualitativa: estudando como as coisas funcionam*. Porto Alegre: Penso, 2011.
- SWAIN, M. Languaging, agency and collaboration in advanced language proficiency. In: H. B. (org.) *Advanced Language Learning: The Contribution of Halliday and Vygotsky* New York: Continuum, 2006.
- _____; KINNEAR, P.; STEINMAN, L. *Sociocultural Theory in Second Language Education: An Introduction Through Narratives*. 2nd ed. Bristol: Multilingual Matters, 2015.
- TAYLOR, L. (Ed.) *Studies in Language Testing 30*. Cambridge: UCLES/Cambridge University Press, 2011.
- TOMASELLO, M. *Origens Culturais da Aquisição do Conhecimento Humano*. São Paulo: WDM - POD., 2003.
- USHAKOVA, T. N. Inner speech and second language acquisition: An experimental—theoretical approach. In: LANTOLF, J. P.; APPEL, G. (ORG.) *Vygotskian approaches to second language research*. Norwood, NJ: Ablex. 1994.
- VÄHÄPÄSSI, A. On the specification domain of school writing. In: PURVES, A.C; TAKALA, S. *An International Perspective on the Evaluation of Written Composition*. Oxford: Pergamon, 1982.
- VAN EK, J.. TRIM, J. *Threshold 1990*. Cambridge: Cambridge University Press, 1998.
- VAN EK, J.. TRIM, J. *Waystage 1990*. Cambridge: Cambridge University Press, 1998.
- VAN EK, J.. TRIM, J. *Vantage*. Cambridge: Cambridge University Press, 2001.
- VAN DIJK, T. A. *Discurso e contexto: uma abordagem sociocognitiva*. Tradução: Rodolfo Ilari. São Paulo: Contexto, 2012.
- VAN LIER, L. Reeling, wrthing, drawling, stretching and fainting in coils: oral proficiency interviews as conversations. In: *TESOL Quarterly*, v. 23, pp. 480-508, 1989.
- VOLOCHINOV, V. *Marxismo e filosofia da linguagem: problemas fundamentais do método sociológico na ciência da linguagem*. 2ª edição. São Paulo: Editora 34, 2018.

VYGOTSKY, L. *Pensamento e linguagem*. São Paulo: Martins Fontes, 1987.

_____. *Thought and Language*. Massachusetts: The MIT Press, 1986.

WEIR, C. O'S. B.; TOMOKO, H. *Exploring difficulty in Speaking tasks: an intra-task perspective*. British Council, Cambridge English Language Assessment and IDP: IELTS Australia. IELTS Research Reports, Volume 6, 2003 Disponível em: <<https://www.ielts.org/teaching-and-research/research-reports>>. Acesso em: jun. 2020.

_____. *Language testing and validation: an evidence-based approach*. Hampshire: Palgrave Macmillan, 2005.

WERTSCH, J. C. *Vygotsky and the social formation of the mind*. Massachusetts: Harvard University Press, 1985.

WHITE, J. *IELTS SPEAKING STRATEGIES*. Publicado por Julia White, 2020

YOUNG, S.; MILANOVIC, M. *Discourse Variation in oral proficiency interviews* in *Studies in Second Language Acquisition* 14, 1992