

Avaliação diagnóstica em contextos migrantes: medidas e formatos de instrumentos nas escolas portuguesas.

Sandra Figueiredo¹, Margarida Martins¹, Carlos Silva²

1 Instituto Superior de Psicologia Aplicada

2 Universidade de Aveiro

Sandra Figueiredo, Instituto Superior de Psicologia Aplicada, Lisboa, sandradfigueiredo@ua.pt

A investigação na área de língua segunda determinou a necessidade de procedimentos de medição sobre a proficiência e competência. Na primeira fase deste projeto científico, pretende-se aplicar e validar provas selecionadas do corpus de testes internacionais, em alunos locutores de Português Língua não Materna do 1.º, 2.º e 3.º Ciclos, de Lisboa e Aveiro. As provas adaptadas apresentam uma proposta alternativa e complementar aos dispositivos do Ministério da Educação, focando na emergência de instrumentos validados. Algumas provas serão discutidas no sentido da complexidade e singularidade da estrutura de diagnóstico em contexto de avaliação de alunos com experiência migratória.

Palavras-chave: avaliação educacional; língua segunda; experiência migratória.

A INVESTIGAÇÃO INTERNACIONAL, O CENÁRIO NACIONAL E MODELOS DE RESPOSTA

Nas últimas décadas a investigação na área de aquisição de linguagem tem focado a área de desenvolvimento de língua não materna, examinando os processos de aprendizagem e produzindo conhecimento acerca dos mecanismos utilizados pelos aprendentes e as suas implicações (Figueiredo, 2012, 2010; 2009a,b,c,d; Gullberg & Indefrey, 2006; Calvo, 2003; Doughty & Long, 2003; Hyltenstam & Abrahamsson, 2003). A investigação internacional na área de aquisição de Língua Segunda propiciou informação válida para a intervenção no ensino e aprendizagem. No entanto, primeiro, os instrumentos e resultados dos estudos teóricos e práticos desenvolvidos não se encontram completamente disponibilizados e clarificados para o contexto português ("open and free access", política FCT) o que fragiliza a intervenção escolar, no contexto particular de Português Língua não Materna (PLNM). Segundo, os instrumentos válidos são raros nos contextos educativo e científico nacionais. A formação educacional da população imigrante é reportada em estudos (Semedo, 2011; Silva & Gonçalves, 2011; Osório & Meyer, 2009; Estudos do Observatório de Imigração (OI) 2005-2011) essencialmente descritivos (embora necessários), documentos orientadores (DGIDC; Conselho da Europa, 2001) ou compilações estatísticas (Alto Comissariado para a Imigração e Diálogo Intercultural, ACIDI, e Instituto de Linguística Teórica e Computacional), mas sem referência a práticas efetivas e instrumentos validados.

Estes são fundamentos que também justificam resultados da Eurostat (2011) relativamente ao grave abandono escolar dos alunos portugueses. Uma das causas é o diminuto conhecimento dos atores educativos acerca de medidas práticas dirigidas a minorias linguísticas (Comissão Europeia, 2009). Os dados e relatórios do OI (unidade do ACIDI) ratificam as necessidades educacionais que as escolas portuguesas ainda não resolveram desde 2005, exortando-se à

investigação científica (OI, 2011; 2012) mais trabalho e evidência no sentido de desenvolver resposta eficaz e métodos para o apoio escolar em PLN, nos planos de desempenho académico e inclusão social. Essa evidência apenas tem sido resultado de investigação internacional (Carrillo, 2011; Rosemarie et al., 2012; Levin & Shohamy, 2008). A OCDE (2011) sublinha a necessidade de qualificação das populações e equidade no acesso à Educação, sendo que os índices escolares portugueses têm sido insatisfatórios na comparação com os pares europeus. Esses índices são fundamentados adicionalmente pelo facto da população imigrante se encontrar sobre representada na vias profissionalizantes, apesar das medidas educativas ensaiadas no triénio 2007-2009 (Seabra et al. (ACIDI), 2011) para encorajar o acesso a outros níveis de ensino. Os alunos “com origem na imigração” são prioridade do II Plano para a Integração de Imigrantes (2010-2013). Os alvos consistem na organização de conhecimento válido da população escolar e a formação dos professores para o ensino de PLN (Seabra et al., 2011; Silva & Gonçalves, 2011).

Noutra perspetiva adicional, a investigação em língua não materna foi restringida à análise de metodologias e técnicas aplicadas pelos professores de línguas estrangeiras, abundando informação sólida mas que não é transferível para o campo de metas de ensino, aprendizagem e avaliação em contexto de L2. No entanto, a questão da aquisição L2, sobretudo no domínio do Português Europeu, ainda é um tema novo (daí que incidimos, na componente prática do programa de trabalhos, na bibliografia internacional e orientações nacionais e europeias) carecendo de análise e de desenvolvimento de programas de avaliação e orientação no contexto escolar. A falta de supervisão sobre a escolha e uso de instrumentos conduz a programas de apoio de PLN ineficientes e que são decalcados a partir das metas de aprendizagem definidas no âmbito das línguas maternas (Pascoal et al., 2006; Direção-Geral de educação, 2012). Por outro lado, as medidas educacionais previstas não têm sido consideradas pois no relatório de monitorização da integração de imigrantes (Bijl & Verweij, 2012) a coordenação portuguesa nunca refere (pp. 291-311) o apoio escolar (ou outra semântica relacionada) como prioridade nem reporta resultados ou métodos didáticos aplicados. Este trabalho apoiar-se-á nos padrões e recursos disponíveis na literatura internacional tais como os seguintes modelos: *International Language Testing Association*, *Repository Alaska Department of Education*, *Minnesota Language Proficiency Assessments*, *California ELD Standard*, *ALBERTA EDUCATION*, *Center for Advanced Research on Language Acquisition*, *Test of English as a Foreign Language*, *TESOL Standards*, principalmente.

BENCHMARKS, VERIFICAÇÃO E VALIDAÇÃO

A validação de instrumentos no âmbito específico de avaliação de competência de alunos aprendentes de língua não materna deverá ter em consideração os seguintes passos que subjazem à organização deste projeto:

- Análise de “benchmarks”: análises de modelos e programas de testes (e estratégias), nacionais e sobretudo internacionais, no mesmo contexto temático e investigativo, sobretudo no domínio de padrões e testes existentes para definir níveis de proficiência e de estimulação) e de tipos de recursos digitais (e de plataformas onde se disponibilizam) internacionalmente.
- seleção de testes que se adequem aos objetivos e intenções de medida: competência, desempenho em conteúdos específicos, proficiência (“nivelamento near-native”) e “placement” (identificação do nível para escolha do método de ensino).
- Distinção dos testes de acordo com os grupos de locutores/minorias e análise da lógica de adequação aos locutores imigrantes portugueses.

No que respeita a este último ponto, conduzimos uma pesquisa intensiva de recursos educativos (digitais sobretudo) existentes e que apresentem produtos direcionados para o Português Língua Não Materna (PLNM). No contexto dos recursos educativos disponibilizados e disseminados pelo Ministério da Educação (e Direção-Geral para a Inovação e Desenvolvimento Curricular – DGIDC/DGE), apresentam-se cerca de 1500 recursos que abrangem áreas curriculares distintas e com dinamismo verificado, mas não apresentam, com frequência desejável, uma estrutura complexa multimédia nem um caráter interativo (e nem 'iterativo'). No entanto refere-se a validação da maioria desses produtos (de acesso livre) o que significa que alguns (cerca de 500) se encontram não cientificamente supervisionados mas acessíveis aos profissionais e comunidade discente. Mas a validação referida não se encontra explicitada devidamente pois a indicação do Ministério é sobre um processo de avaliação (dos recursos educativos) que se rege por verificação de erros científicos (conforme se prevê, deve aplicar-se ao domínio do conhecimento científico transmitido na proposta dos autores dos recursos), erros linguísticos (domínio da correção dos enunciados presentes nos enunciados); deteção de estereótipos ou enviesamento de conceitos culturais (talvez este seja um critério que deveria transformar-se mais frequentemente em produto educativo e não apenas como critério de avaliação); desrespeito pelos direitos de autor e propriedade intelectual. No entanto este conjunto de indicadores não completa um processo de validação em termos científicos, são formas válidas e gerais de análise mas carecem da aplicação empírica dos recursos educativos em estudos, pelo menos, exploratórios. Daqui depreender-se a exequibilidade da aplicação e utilização de tais produtos em sala de aula e não os excluir de outros contextos em que possam ser considerados de forma 'válida'. Um dos aspetos que deteta a necessidade de supervisão científica de alguns produtos, no contexto de língua não materna, prende-se com a atribuição de categorias descritivas (do produto) de forma indiferenciada para vários recursos. A validação que pretendemos no projeto será direcionada a testes de avaliação, portanto, não englobando prioritariamente, mas não excluindo, produtos educativos de entretenimento para a aprendizagem de língua não materna, em particular. Højsholt-Poulsen (s/d, p. 29) apresenta a questão da validação de forma muito coerente pois afirma que no contexto dos recursos digitais dinamarqueses não há validação, há sim "verificação" dos dados dos produtos e da forma como se apresentam por parte dos seus autores. A verificação seria, no caso português, uma terminologia também mais adequada dado que aferir produtos em termos de validação ainda exige maior envolvimento longitudinal da investigação científica.

CONSTITUIÇÃO E RACIONAL DE UMA BATERIA DE PROVAS DE DIAGNÓSTICO

Num primeiro momento apresentamos a seleção final dos testes considerados para este ensaio diagnóstico em Portugal, na população escolar imigrante. Num segundo momento descrevemos alguns aspetos fulcrais que constituem o racional da organização do conjunto de testes nesta investigação. Com base nos testes disponíveis na literatura e observando os materiais facultados a partir do contacto com as respetivas entidades e autores, na área de avaliação diagnóstica de competências e proficiência em L2, seleccionámos algumas provas específicas para parcial ou integral adaptação, que de seguida se identificam.

FASE I - DOCENTES

a) Inventory of undergraduate and graduate level reading, writing, speaking, and listening tasks educational testing service (Educational Testing Service, Rosenfeld, Leung, & Oltman (2001, pp. 63-70). A análise dos resultados será de acordo com várias categorias: localização de informação, compreensão básica, aprendizagem, integração, conteúdo, organização, desenvolvimento, linguagem, factos e detalhes, vocabulário, ideias principais, inferências, funções comunicativas. Estas categorias estão associadas às competências de leitura, escrita, oralidade e compreensão oral. O questionário a aplicar em contexto de escolas portuguesas será

baseado em princípios estruturais desse inquérito, sendo que se adaptarão a maioria dos enunciados. Neste momento, a aplicação já teve início em escolas da área de Lisboa.

b) Teacher Post-Rating Questionnaire (“Issues and Arguments in the Measurement of Second Language Pronunciation”, Isaacs, 2010, p. 226). Questões relacionadas com a compreensibilidade e inteligibilidade de discurso serão analisadas pelos professores das escolas.

Neste momento, a aplicação destas duas partes do questionário de docentes já teve início em escolas da área de Lisboa.

FASE II - ALUNOS

- Teste lexical com base no “Lextutor” (listas de palavras organizadas por níveis de frequência e de regularidade, em Inglês, e de acordo com a respectiva definição semântica. Este teste torna-se mais válido que os comuns testes de vocabulário sim/não, pois testa a compreensão e não apenas a decodificação, Alderson & Banerjee, 2002, p. 90) de Paul Nation (1990) e com apoio dos indicadores providenciados pelo Corpus Lexicográfico do Português (CLUL), especificamente o CORLEX que apresenta a lista lexical portuguesa (Português Europeu) de acordo com a frequência (critério adotado: palavras registadas com mais de 1000 ocorrências no idioma). As palavras serão escolhidas atendendo a esse fator, o qual costuma ser negligenciado nos testes de avaliação com população migrante. O lextutor providencia, por outro lado, o formato aproximado do desejado neste projeto pelas características abordadas. Numa segunda tarefa dentro do teste lexical teremos uma prova exploratória com o objectivo de observar a frequência necessária de vocábulos para a sua aquisição por parte do locutor estrangeiro. Em três textos encontram-se palavras repetidas, distribuídas por três núcleos semânticos (janelas, galáxia, cobras) de forma a avaliar a captação desses vocábulos por parte do locutor, em correlação com o índice de frequência definido. Os textos foram analisados na medida LEXILE (espanhol) o que informou uma medida de dificuldade dos textos aproximada em Português: cada texto varia entre 610 L e 1000L. Os textos foram escolhidos e preparados de forma a que correspondam aos níveis esperados para o 3.º ano do 1.º Ciclo e o 8.º e 9.º anos, segundo os padrões do LEXILE (Lexile-to-Grade Correspondance).

- Teste de extracção morfológica com base no “Extract the Base Test” (teste de alterações parciais ou simultâneas em termos fonológicos e ortográficos, em línguas inglesa e espanhola) de August e Kenyon (2001), sendo que será adaptado o exercício apenas com base na língua segunda (Português) e não atendendo à língua materna dos alunos.

- Teste de palavras cognatas com base no “Cognate awareness test” (testes de palavras cognatas entre a língua materna e a língua segunda, nos casos específicos do Inglês e do Espanhol) de August, Kenyon, Malabonga, Louguit, Caglarcan, e Carlo, M. (2008).

- Teste de transferência fonológica com base no “Phonological Transfer Test” (teste de avaliação da transferência e interferência fonológica entre o inglês e o espanhol, com exercícios de segmentação de pseudopalavras) de Rolla e August (2002).

Estes três últimos testes foram originalmente desenvolvidos no contexto de um programa de investigação sobre competências de transferência e leitura de crianças bilingues. Serão considerados os testes no contexto específico de língua segunda e com várias idades.

- Teste de escrita (narrativa e reconto) com base em sequência de imagens e nos descritores das provas de escrita e reconto de ALBERTA. CA. No que respeita à parte de reconto, será seguido o tipo de registo do teste português de diagnóstico, acedido livremente no site da Direção Geral de Educação (<http://www.dgidc.min-edu.pt/outrosprojetos/index.php?s=directorio&pid=195>). Será apresentada uma breve narrativa, em forma oral, e será pedido ao aluno a reconto atendendo a: utilização de vocabulário adequado, ordem dos eventos da história, construção de enunciados frásicos com boa sintaxe. Por outro lado, são observadas também as demandas cognitivas de forma a evitar léxico com múltiplos níveis de sentido, estrutura de textos complexa, eventos desordenados, linguagem figurada e ambígua, registo académico, intertextualidade de nível elevado (Common Core State Standards for English Language Arts & literacy in History/Social Studies, Science, and Technical subjects). A verificação da estrutura dos textos e demandas cognitivas também integra o conjunto de indicadores do Center for Advanced Research on Language Acquisition (CARLA). Neste repositório encontramos alguns aspetos importantes também na avaliação de proficiência, embora para alunos mais velhos (Minnesota Language Proficiency Assessments (MLPA)).

- Os exercícios implicados na estrutura da bateria Woodcock-Munoz Language Survey-Revised (WMLS-R) serão considerados tais como a nomeação com base em imagens (“expressive language semantic”), associação de palavras (sinonímia e antonímia) e analogias verbais, sendo que estes testes encontram-se também no “teste de proficiência” (cluster assim denominado pelos autores da bateria Bilingual Verbal Ability), de Munoz-Sandoval, Cummins, Alvarado, & Rued, M. (1998). Bilingual Verbal Ability Tests Comprehensive Manual. Itasca, IL: Riverside Publishing. As imagens utilizadas especificamente na prova de nomeação oral e escrita são oriundas do exercício de nomeação escrita do Teste Diagnóstico da DGE (elaborado pelo ILTEC, 2009), já elaborados para a avaliação de competências em PLNM.

- teste de compreensão oral e leitura do conjunto Test Of English as a Foreign Language (TOEFL). Os textos, em português, a serem propostos para leitura e audição serão escolhidos de acordo com critérios do LEXILE, atendendo à equação de dois factores: dimensão das frases e das palavras e frequência lexical. A idade e a qualidade do texto não se encontram em análise daí que o Lexile foca normas que importam substancialmente na área de avaliação em L2. Nesta área a idade também não é o fator que deve única e diretamente corresponder com a proficiência e daí que as demandas cognitivas das metas académicas para cada ano escolar não correspondem sempre à idade do aluno em condição de experiência migratória. No entanto, o texto seleccionado apresenta uma medida (680L, inferida por aproximação com as normas de cálculo com a medida espanhola) que se situa no intervalo determinado pelo LEXILE para o 4.º ano de escolaridade: 445L-810L. Como se tratam de alunos em condição não nativa, o texto escolhido (para ser lido e ouvido) será aplicado a anos escolares seguintes também.

Racional dos testes na FASE II: alguns princípios da estrutura

Testes dirigidos, de forma geral, a 4 competências: compreensão oral, pronúncia e organização de texto (na oralidade), fluência (na compreensão escrita/leitura) e

edição/composição (na escrita) (Alberta, 2010, p. 4; TOEFL, 2011, p. 9). No entanto a produção oral será a competência não visada no teste diagnóstico em elaboração. No teste do ILTEC as várias provas incluem essa competência, contudo neste projeto as 9 provas desenvolvidas apresentam uma estrutura diferenciada e que não incluiu a oralidade. O princípio orientador do teste segue exercícios já testados neste tipo de população, com base em provas internacionais. Adicionalmente, com base na estrutura dos testes da TOEFL (2011, p. 5), com última aplicação conhecida em Maio de 2012, os testes poderão apresentar este formato:

- secção do teste de oralidade e de compreensão oral (exposição oral e leitura, os textos incidem em linguagem e temas de cariz académico, portanto de acordo com os conteúdos curriculares): nesta secção serão focados o vocabulário e expressões idiomáticas (TOEFL, 2011, p. 9), construção sintática de acordo com as fórmulas mais correntes na linguagem oral portuguesa, textos breves com estilos próximos dos géneros textuais em utilização académica, de acordo com os níveis de escolaridade (nos relatórios da TOEFL encontramos aplicações sobretudo em população adulta, incidindo nas universidades, sobretudo no que respeita à avaliação conduzida sobre a reconstrução dos testes, Rosenfeld et al., 2001). O material a ser utilizado (texto) seguirá os critérios normativos de frequência lexical (CORLEX), usados também noutros testes em contexto de avaliação em língua materna. Nesta secção vários elementos estão a ser implicitamente testados: competência linguística, discursiva, funcional e sociolinguística. A pronúncia não constitui alvo de avaliação na medida em que os testes utilizados, na verdade, não estão normalmente desenvolvidos enquanto testes válidos para determinar sobre a acentuação e características de sotaque (ou ausência) dos sujeitos (Isaacs, 2010).

- secção do teste de escrita e compreensão escrita (de acordo com o critério de linguagem académica): nesta secção serão focados a estrutura gramática mas focando os aspetos ao nível do fonema, em Português, com base em temas gerais e não cientificamente específicos (sem destacar áreas curriculares dos programas). No caso do tipo de textos apresentados, a familiarização com os textos pode interromper a compreensão estrutural e gramatical dos textos em si (Alderson & Banergee, 2002) na medida em que gera oportunidades (indesejadas em avaliação) de resposta adivinhada sem informação compreendida. Quer na escrita, quer na oralidade, a literacia também é uma das competências a ser avaliadas (por exemplo através de um sumário de excertos textuais ouvidos ou lidos) (Hudson, 1996, p. 12). Na compreensão escrita, as questões deverão ser diferenciadas em três sentidos: questões que apelam à informação explícita num texto de partida; questões que apelam à informação implícita que deve ser inferida; questões que são 'encriptadas' e apenas podem ser respondidas com background linguístico (Alderson & Banergee, 2002, p. 85). Neste plano dual, distingue-se dois níveis de competência a priori.

Serão analisadas várias categorias (Rosenfeld et al. 2001, pp. 18-22) para cada uma das quatro competências em causa. Desta forma:

- discriminação auditiva: factos e pormenores; argumentos principais do texto; vocabulário e terminologias; inferências (a interpretação).

- leitura: compreensão de nível complexo (com presença e ausência do suporte texto); compreensão básica; integração (síntese de textos ou ideias de um texto). Este nível em princípio será desconsiderado na bateria prevista para aplicação nesta investigação.

- escrita: organização; linguagem (domínios das estruturas da língua); conteúdo; argumentação; diversidade lexical.

- oralidade: fluência; organização de texto; deixis.

Na temática dos testes no âmbito da competência escrita, Hamp-Lyons e Kroll (1997, p. 11) consideraram que “Without an understanding of the nature of academic writing tasks, it would be impossible to construct a test”. Na estrutura do teste de escrita ter-se-á em conta, de acordo com os autores referidos:

“1 locus of writing, i.e., in class or out of class 2. length of product to be written 3. genre of product 4. cognitive demands of task 5. rhetorical properties of product, subdivided into mode of discourse and specification of rhetorical patterns” (p. 12). Os textos serão construídos, além da consideração dos tipos textuais e das temáticas, de acordo com diferentes esquemas linguísticos de forma a captar a proficiência real dos indivíduos, já que textos com estruturas gramaticais internamente variadas testam mais significativamente o poder de ajustamento e de interpretação em termos gramaticais:

types of text, one in which the schema activated by the first part of the text is not congruent with the subsequent linguistic input (‘non-matching’) and the other in which the schema is uniform throughout (‘matching’). They argue that ‘non-matching’ texts demand that test takers process the incoming linguistic cues quickly and accurately, adjusting their schema when necessary. (Alderson & Banerjee, 2002, p. 89).

Na análise de conteúdo serão considerados os critérios do Center for Advanced Research on Language Acquisition e para o registo das observações e classificações dos desempenhos serão usados as folhas de registo de Alberta Education – ESL Proficiency Benchmarks. Ainda, teremos em conta, de forma complementar, teste de consistência ortográfica (relação de transparência e opacidade fonética; e regularidade), teste de fluência e pronúncia (Isaacs, 2010), e testes de coesão e coerência textual (DGIDC, 2009). De acordo com Abeywickrama (2007), a coesão e coerência textual no contexto de L2 é testada de forma válida através de dois tipos de testes que consideraremos na nossa bateria: composição escrita e exercício de completamento de enunciados. O domínio da coesão e coerência são considerados preditores (Abeywickrama, 2007) da compreensão da leitura em L2. Por outro lado, a compreensão e inteligibilidade são dois conceitos a ser analisados. Quanto à inteligibilidade, a densidade lexical é uma das formas que permite a análise do construto,

the terms “intelligibility” and “comprehensibility” have not been consistently applied across studies, leading to definitional confusion and difficulties in making cross-study comparisons (Isaacs, 2008). Derwing and Munro (1997) provide useful and conceptually clear definitions of these terms, which have been adopted by several researchers (e.g., Kennedy & Trofimovich, 2008) and are also adhered to in this thesis. Intelligibility denotes the amount of the L2 learner’s utterance that the listener is able to understand, and is most often measured by the proportion of words that the listener is accurately able to transcribe, although other methods, such as true/false sentences (Derwing, Munro, & Wiebe, 1997) and comprehension questions (Hahn, 2004) have also been used. (Isaacs, 2010, p. 17).

Quanto à 'compreensibilidade' ("comprehensibility"), a análise é distinta: "defined as listeners' perceptions of ease of understanding of L2 speech and is most often measured on a numerical (Likert-type) scale punctuated with the scalar endpoints "extremely easy to understand" and "extremely difficult to understand"(Isaacs, 2010, p. 18). Nesta linha de ideias, em avaliação teremos inteligibilidade (quantidade e tipo de léxico), a compreensibilidade (percepção e entendimento, em termos de qualidade, do discurso em L2), fluência (relacionado com ritmo e tipo de cadência do discurso), pragmática e prosódia. Liscombe (2009) interessou-se pela capacidade dos sujeitos perceberem os atos ilocutórios (nas suas características emocionais, portanto ao nível paralinguístico, e pragmáticas, através do questionamento da informação contida nos textos, por exemplo ao nível específico da coesão textual) dos emissores dos textos com base nesses aspectos.

A construção da bateria encontra-se definida e o estudo exploratório em curso, em escolas do distrito de Lisboa. Pretende-se verificar a exequibilidade dos testes adaptados e a sua validade científica.

REFERÊNCIAS BIBLIOGRÁFICAS

Abeywickrama, P. (2007). *Measuring the knowledge of textual cohesion and coherence in learners of English as a second language (ESL)*. Tese de Doutoramento. University of Hong Kong.

Alderson, J., & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching*, 35, 79-113.

California ELD Standards (1999). The English Language Development Standards for the State of California and the IPT Battery of Tests for English Language Learners <http://www.ballard-tighe.com/pdfs/alignmentAssessment/ipt-tesol2006.pdf>. Ballard & Tighe.

Center for Advanced Research on Language Acquisition (CARLA). Contextualized Reading Assessment (CoRA). Minneapolis.

Center for Advanced Research on Language Acquisition (CARLA). Developing Speaking and Writing Tasks. Minnesota Language Proficiency Assessments (MLPA).

Common Core State Standards Initiative, Preparing America's Students for College & Progress. Common Core State Standards for English Language Arts & literacy in History/Social Studies, Science, and Technical subjects. http://www.corestandards.org/assets/CCSSI_ELA%20Standards.pdf

Direção-Geral da Educação (2012). Teste de Diagnóstico de Português Língua Não Materna (1.º ano – 12.º anos). <http://www.dgidec.min-edu.pt/outrosprojetos/index.php?s=directorio&pid=195>.

Figueiredo, S. & Silva, C. (2010). Factores afectivos e desempenho cognitivo na aprendizagem linguística. Tese de Doutoramento. Universidade de Aveiro.

Figueiredo, S., & Silva, C. (2009a). Tests battery of cognitive achievement in second language context. Teste electrónico. Universidade de Aveiro: Departamento de Ciências de Educação e Centro Multimédia de Ensino à Distância.

Figueiredo, S., Silva, C. (2009b). Comportamento textual e discriminação de identidades articulatórias: desempenho diferenciado em aprendentes de Português Língua Segunda. *Revista de Avaliação Psicológica*, 8 (3), 289-301. ISSN: 1677-0471. [F.I. 0.2295] .

Figueiredo, S., & Silva, C. (2009c). Cognitive differences in second language learners and the critical period effects. *L1 – Educational Studies in Language and Literature*, 9(4), 157-178.

Figueiredo, S., Silva, C. (2009d). Language skills assessment: electronic test application with second language learners. *The International Journal of Learning*, 6(8), 559-578.

Hamp-Lyons, L., & Kroll, B. (1997). *TOEFL 2000: Writing: Composition, Community, and Assessment*.

http://www.ets.org/research/policy_research_reports/publications/report/1997/iciy.

Hudson, T. (1996). *Assessing Second Language Academic Reading from a Communicative Competence Perspective: Relevance for TOEFL 2000*. ETS Research Memorandum. http://www.ets.org/research/policy_research_reports/publications/report/1996/hzcg.

Isaacs, T. (2010). *Issues and Arguments in the Measurement of Second Language Pronunciation*. Tese de Doutorado. McGill University, Montreal.

Liscombe, J. (2007). *Prosody and Speaker State: Paralinguistics, Pragmatics, and Proficiency*. Tese de Doutorado. Columbia University.

Minnesota Language Proficiency Assessment (MLPA). <http://www.carla.umn.edu/assessment/MLPA.html>. Minnesota.

Pascoal, J. & Oliveira, T. (2005). *Português Língua Não Materna no Currículo Nacional: Orientações Nacionais: Diagnóstico de competências em Língua Portuguesa da População escolar que frequenta as escolas portuguesas*. CAPLE (Centro de Avaliação de Português Língua Estrangeira), Faculdade de Letras da Universidade de Lisboa. www.min-edu.pt/np3content/?newsId=61&fileName=plnm_orient_nacionais_diagnostico.pdf.

TOEFL (2012). *TOEFL® Institutional Testing Program (ITP) & TOEFL Family Products*. Brasil.

<http://www.ets.org/pt/toefl/>

TOEFL (2011). *Test and Score Data Summary for TOEFL iBT® Tests and TOEFL® PBT Tests*. Educational Testing Service

Woodcock, R. W., Munoz-Sandoval, A. F., Rief, M.L. & Alvarado, C. G. (2005). *Woodcock-Munoz Language Survey – Revised, English*. Itasca, IL: Riverside Publishing.