

# **INTEGRAÇÃO E VALIDAÇÃO DE TESTES PIONEIROS EM CONTEXTO ESCOLAR MIGRANTE: PERSPETIVAS SOBRE AVALIAÇÃO DIAGNÓSTICA E DESEMPENHO ACADÉMICO EM LÍNGUA SEGUNDA.**

**Sandra Figueiredo<sup>1</sup>, Margarida Martins<sup>2</sup>, Carlos Fernandes da Silva**  
**1Instituto Superior de Psicologia Aplicada (Portugal)**

**2Instituto Superior de Psicologia Aplicada (Portugal)**

**3Departamento de Educação, Universidade de Aveiro (Portugal)**

**sandradfigueiredo@ua.pt**

## **Resumo**

Nas últimas décadas a investigação científica na área de aquisição de linguagem tem focado a dimensão de desenvolvimento de segunda linguagem, examinando os processos de aprendizagem dos sujeitos e produzindo conhecimento acerca dos mecanismos utilizados pelos aprendentes e as suas implicações em termos de desempenho académico. A investigação internacional na área de aquisição de Língua Segunda (L2) propiciou informação válida para a intervenção no ensino e aprendizagem. No entanto, os instrumentos e resultados dos estudos teóricos e práticos desenvolvidos, em contexto português, não se encontram disponibilizados o que impede a intervenção nas escolas, no contexto particular de Português Língua não Materna (PLNM). Ainda, verifica-se uma ausência preocupante de instrumentos válidos, nos contextos educativo e científico. Estes fatores conduziram a um projeto de pós-doutoramento que tem como objectivo a construção de uma bateria de avaliação diagnóstica em contexto de PLNM. Desta forma encontra-se em curso um plano organizado por quatro fases: seleção dos instrumentos internacionais (e nacionais existentes) para adaptação ao contexto de Português como Língua Segunda (L2); validação dos instrumentos na versão portuguesa; desenvolvimento de um repositório eletrónico com o objetivo de divulgar as medidas selecionadas, adaptadas e/ou desenvolvidas (após processo de verificação de propriedades psicométricas); e disseminação dos resultados e do repositório. Neste trabalho serão apresentados dados preliminares descritivos relativamente à seleção e tradução dos testes, argumentando sobre os racionais construídos para a estrutura dos testes e respectivas formas de pontuação. A seleção dos instrumentos encontra-se focada em provas da literatura sobretudo internacional, com propriedades psicométricas válidas. Os testes escolhidos incidirão nas competências de oralidade e escrita (compreensão, produção) e providenciarão informação adequada para os contextos escolares reais da população migrante. Paralelamente será apresentada uma breve revisão crítica dos instrumentos nacionais disponíveis e do seu estado de validação, com aplicação nas dimensões social e cognitiva da aprendizagem de uma L2 em contexto formal e educacional.

**Palavras-chave:** avaliação em Língua Segunda, desempenho académico, alunos imigrantes, validação.

In the last decades the scientific research in language acquisition area focused the second language dimension, examining the learning processes of individuals and providing knowledge regarding the learner's mechanisms and their implications in academic achievement context. The international research in Second Language Acquisition (SLA) provided valid information for intervention in teaching and learning. However, instruments and data from theoretical and practical studies, in portuguese context, are unavailable which difficult intervention in schools, in the specific context of Portuguese as Second Language. Additionally the absence of valid

instruments, in educational and scientific settings, is of great concern. These factors motivate a post-doctoral project with the main goal on the construction of a diagnostic tests battery in Portuguese SL context. At this moment the project is scheduled through four phases: selection of international instruments (and national tests available) to adapt for the portuguese SL context; validation of the instruments in portuguese version; development of an electronic repository regarding the dissemination of the selected, adapted and developed tests (after the verification of validation process); and dissemination of the results and of the repository. In this paper will be presented preliminary descriptive data regarding the selection and translation processes, arguing on the rationales that underlie the structure and scores of the tests. Selection of tools is centred in the main international research, with valid psychometric properties. The selected tests will have focus on orality and writing skills (comprehension and production abilities) and will provide evidence for real migrant school contexts. Alongside a critical review of instruments will take place, regarding its validation and application, attending to the social and cognitive dimensions involved in the SL context learning and formal setting.

Key-words: assessment in second language, academic achievement, immigrant school students, validation.

## 1 INTRODUÇÃO

Um dos aspetos que caracteriza o produto último desta investigação será, na área da aprendizagem de língua não materna, que o trabalho de validação de instrumentos de avaliação se converta em recurso dinâmico (portanto considerando que os processos de validação de instrumentos são morosos e implicam também disponibilidade de circunstâncias da população escolar) num futuro repositório de testes e recursos. Essa 'conversão' categorizará a função que mais nos interessa: estrutura-satélite, ou seja, um instrumento de recursos em permanente vínculo de orientação com as escolas do país e com alguma distinção por região. Para que a estrutura-satélite funcione prevemos um plano de tarefas para constituir um programa de preparação dos diretores escolares (focando uma das prioridades da OCDE, Pont et al., 2008a; 2008b: liderança e sucesso académico) para:

- conhecimento dos instrumentos validados neste projeto e os seus contextos de avaliação;
- conhecimento dos instrumentos anteriores a este projeto a sua aplicabilidade;
- critérios de seleção dos instrumentos de acordo com as populações escolares;
- conhecimento dos resultados expectáveis;
- conhecimento dos procedimentos de aplicação e de interpretação dos resultados;
- organização dos existentes programas de apoio aos alunos em causa, de acordo com os resultados.

A formação pretende-se numa larga escala à semelhança do projeto da Microsoft e Ministério da Ciência e Tecnologia na formação de diretores escolares em Maio de 2012, no âmbito específico de gestão e administração escolar. Contudo a formação para utilização dos instrumentos a ser validados também se destina aos professores e educadores, dos vários níveis de ensino. Neste trabalho de comunicação, portanto de envolvimento, com as escolas interessa-nos que a disseminação envolva os docentes na opinião e sugestão de recursos, de transformação de testes de acordo com contextos que só os profissionais conhecem em especificidade. O recurso à técnica de grupos focais poderá ser pertinente e localizada de forma a ter uma visão mais próxima do sucesso da implementação dos novos recursos que este repositório pretende criar. Desta forma também os docentes poderão integrar a decisão de recursos dirigidos a esta específica população de alunos. Por outro lado, a existência de testes dirigidos a diferentes minorias irá diminuir a distância entre culturas e estereótipos que não é trabalhada nos manuais escolares de língua materna (Observatório de Recursos Educativos). Um efeito pretendido é a parceria do novo repositório com outros similares, nesta área educativa, com repositórios estrangeiros, formulando o projeto adicional de uma espécie de "pan-european forum".

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 e que subjazem a este 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), apresentam-se cerca de 1500 recursos que abrangem áreas curriculares distintas e com dinamismo verificado, mas não apresentam muitas vezes uma estrutura complexa multimédia nem um carácter interativo. 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 deste Repositório 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.

No que respeita a repositórios de recursos digitais, entendendo aqui os repositórios educativos, é curiosa a reflexão de Højsholt-Poulsen (s/d, p. 22) acerca dos conteúdos que devem ter os recursos digitais e que, assim, tornem os repositórios realmente úteis: “devem incluir conteúdos que ajudem o principiante a começar, porque o início é muito importante para que os professores adotem os recursos digitais. E devem ser abertos, de forma a que os professores inovadores possam utilizá-los com os seus próprios conteúdos.”. Outras considerações do autor importam-nos neste Projeto: o acesso dos produtos por parte dos ‘utilizadores finais’ e a abordagem radical da Coreia ou Hungria que apresenta os seus recursos educativos (digitais) num só local e para todos os níveis de ensino. Este último ponto coloca em foco a sinergia difícil de acontecer em Portugal pois há muitos recursos, nem todos válidos para determinados contextos e competências, mas não há encontro entre plataformas e mesmo entre comunidades de investigação que desenvolvem ideias e produtos nas mesmas áreas de Educação. Um problema adicional se desenvolve: desenvolver recursos com apoio dos ministérios ou das entidades de ciência e ensino superior. Neste contexto responde a Investigação, no entanto a disseminação é responsabilidade de todos os agentes envolvidos. O

resultado mais frequente é que os utilizadores não têm conhecimento dos recursos, daí que os repositórios são as soluções atuais mais previstas na comunicação entre os autores e os utilizadores de recursos (Hojsholt-Poulsen, s/d, p. 28), sejam de entretenimento, sejam de avaliação. Outro assunto é a adequação temporal, há recursos que devem ser renovados a cada ano letivo ou noutra período de tempo que se considere suficiente para alterar dados nos repositórios. Além disso, tal como o autor referido anteriormente adiantou, a avaliação do estado e da utilização do repositório tem de ser atenta. O Projeto, em língua não materna, terá em consideração todos estes aspetos (validação, acuidade, avaliação, interação e utilização) embora se refira a uma área muito particular e que, por outro lado, os repositórios europeus ainda não focaram cientificamente testes e padrões de avaliação e intervenção dirigidos à população escolar de locutores de línguas maternas que não o Português.

## 2 SWOT: UMA ANÁLISE DOS CONTEXTOS E PERFIS DE RECURSOS EDUCATIVOS

Um dos primeiros procedimentos nesta investigação é compreender como alguns dos principais repositórios e formas de recursos educativos existentes se organizam e que tipo de estrutura, e sobretudo de funcionalidade, oferecem a nível internacional e nacional. Desta forma conduziu-se por um lado a uma sistematização de dados sobre diversas plataformas que se assumem como guias de recursos educativos, e, por outro lado, à identificação de soluções dirigidas à avaliação e intervenção em população com experiência migratória,

- “Portal da Educação”, Brasil ([http://portal.mec.gov.br/index.php?option=com\\_content&view=article&id=13018:guia-de-tecnologias&catid=195:seb-educacao-basica&Itemid=948%20%20](http://portal.mec.gov.br/index.php?option=com_content&view=article&id=13018:guia-de-tecnologias&catid=195:seb-educacao-basica&Itemid=948%20%20)), exibe um guia de tecnologia organizado em vários blocos que visam gestão da Educação, ensino-aprendizagem e formação, entre outros blocos. No entanto, no módulo “Educação Inclusiva”, não se apresentam recursos ou programas ou testes para a população educativa, apesar de se referir um modelo de sensibilização para a diversidade cultural. É descritivo mas demasiado discreto e teórico no que importa à realidade: validações psicométricas, construção de instrumentos, avaliação em campo, repositórios de recursos.
- “National Action Plan on integration”, Alemanha (<http://www.bmi.bund.de/EN/Themen/MigrationIntegration/Integration/NationalActionPlan>), com algumas medidas, pouco descritivas, de programas intensivos de ensino da língua para os alunos jovens imigrantes. Os incentivos do programa germânico centram-se na integração e não na avaliação e intervenção. Integrar é um nível que envolve maior transação, portanto mais apostas de natureza política e social.
- “Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado”, Espanha (<http://www.ite.educacion.es/es/recursos>), apresenta uma série de recursos para professores e alunos. No entanto o ensino de línguas incide nas línguas materna e nas estrangeiras previstas no programa curricular. Os recursos são projetos muito interativos e bons dispositivos complementares em sala de aula de línguas, mas não se registam produtos dirigidos ao caso de língua não materna. A área da interculturalidade é contemplada mas sem qualquer recurso no domínio linguístico, apenas sensibilização quanto à religião e culturas.
- “National School Supply & Equipment Association”, E.U.A. (<http://www.nssea.org/>), apresenta alguns recursos mas apenas um dos programas digitais é dirigido aos aprendentes de Inglês como Língua Segunda (L2).
- “Portail national des professionnels de l'education”, França (<http://eduscol.education.fr/pid26435/enseigner-avec-le-numerique.html>), os recursos dispostos não visam a aprendizagem ou facilitação de acesso à língua materna por parte de imigrantes e locutores de minorias linguísticas.
- “The Swedish National Agency for Education”, Suécia ([http://www.skolverket.se/om-skolverket/in\\_english](http://www.skolverket.se/om-skolverket/in_english)), várias medidas e planos de ação são dispostos nesta agência de educação e dirigidos a todos os níveis de ensino (exceto o ensino superior) e visando várias populações escolares mas não se mencionam os grupos migrantes.
- “National Digital Learning Resources”, Irlanda (<http://www.ndlr.ie/>), apresenta vários módulos e programas digitais direcionados ao ensino e aprendizagem mas sem alvos na população migrante e suas necessidades educativas. A política do repositório em causa é estimulante e interativa mas carece de produtos na área referida.

- “Open Educational Resources” (<http://www.oercommons.org/>), um repositório dinâmico e de navegação muito fácil que apresenta listas de produtos e bibliografia em várias áreas educativas.

- “Techonology, inovation, education” (<http://www.gesci.org/index.html>), website sobre orientações, programas e recursos com vista ao desenvolvimento da literacia digital e das competências sobretudo nos países mais subdesenvolvidos. De destacar, no que nos interessa, algumas medidas que se intentam face a “Science, Mathematics and English e-resources developed and accessible to students and teachers”.

Um dos problemas observados na análise de vantagens de acesso a recursos educativos deveu-se ao facto de várias plataformas - Dinamarca, Holanda, Hungria, Itália, Finlândia, Noruega, Islândia, Polónia - apresentarem a informação apenas nas respetivas línguas maternas.

Em Portugal encontramos uma adesão significativa a plataformas digitais com recursos educativos aplicáveis a vários domínios escolares mas que não se dirigem, em destaque, às competências em causa: compreensão, codificação e descodificação em língua não materna,

- Pacotes de recursos digitais mais utilizado em Portugal, mas sem supervisão científica (e criado essencialmente por professores) no repositório Promethean Planet (<http://www1.prometheanplanet.com/pt/>). Não há a opção de língua não materna, referente ao Português, apenas a língua materna e a estrangeira.

- “Recursos Educativos Digitais” (<http://recursoseducativosdigitais.blogspot.pt/p/colecoes.html>), vários recursos são inseridos numa plataforma com um estilo mais informal mas apresentando uma variedade de recursos e de websites úteis à Educação. Todavia não se encontra o domínio dos recursos para o contexto de língua não materna.

- “Plataforma Digital de Apoio escolar” (<http://www.20.leya.com/entrada/plataforma.html>) do Grupo editorial LEYA, um espaço que realça as vantagens, para professores e alunos, das aulas em formato digital. São recursos complementares às aulas tradicionais. Sem destaque para aulas (ou espaços de apoio curricular) de língua não materna.

- Repositório de recursos educativos digitais disponível no “Portal das Escolas” ([https://www.portaldasescolas.pt/portal/server.pt/community/00\\_recursoseducativos/259](https://www.portaldasescolas.pt/portal/server.pt/community/00_recursoseducativos/259)) que apresenta cerca de 1500 dispositivos, de criação variadas e em diversas áreas curriculares, mas nem todos se encontram aferidos. Na questão da validação, já anteriormente referida, alguns critérios são considerados como formas de validar cerca de 1000 programas (nem todos foram sujeitos a supervisão) mas que não correspondem a verdadeiros critérios de validação, sobretudo considerando a aplicação dos recursos em populações escolares (estudos exploratórios). Na análise feita neste Portal detetámos que, no catálogo dos ‘mais vistos’, os temas relacionados com a aprendizagem de língua não materna não eram os principais abordados. Encontramos referências à imagem de interculturalidade (em Formação Cívica), sensibilização para a distinção de raças, os recursos de Língua Portuguesa raramente são destinados aos imigrantes, na prioridade da definição da população alvo. Alguns recursos na área fonética e lexical (Linguística) são explorados no contexto da língua não materna, efetivamente. De entre os recursos mais votados, os conteúdos em língua Portuguesa existem mas sem referência significativa para a Língua não Materna. Os aspetos fonéticos e lexicais são os mais focados, no entanto. Por fim, na pesquisa dos recursos mais recentes, os conteúdos para locutores estrangeiros não surgem. Por outro lado, estes recursos nunca visam a avaliação de competências, proficiência e aspetos afetivos face à língua e contexto não maternos. Os critérios definidores de produto (a ficha técnica dos recursos), no que respeita a língua não materna, apresentam-se simplistas e assim definidos: densidade semântica média, dificuldade média e quase sempre para todos os anos de escolaridade. Os níveis de proficiência são, na maioria, A1 e A2 (Comissão Europeia, 2001). A indiferenciação dos descritores em todos os recursos é um problema.

## **2.1 Pressupostos teóricos e práticos da estruturação do teste diagnóstico**

O teste diagnóstico em construção baseia-se em duas fontes principais assim determinadas neste específico estudo:

- padrões ('benchmarks') definidos por ano de escolaridade e por nível de proficiência (sobretudo adotando a perspectiva americana) e provas de avaliação disponíveis em repositórios estrangeiros relacionados com a educação de alunos imigrantes (ALBERTA, 2010;2012), que providenciam informação e dados válidos sobre a avaliação a conduzir em contexto escolar;
- a evidência científica que os quadros teóricos desenvolvem no que respeita ao entendimento dos processos de aquisição de L2.

No que se refere ao primeiro ponto, após um estudo analítico de padrões e formas de avaliação disponíveis em repositórios internacionais, constatamos vários modelos consistentes sobre a avaliação e intervenção em contexto de língua segunda ou estrangeira. Seleccionámos alguns desses modelos a partir dos quais baseamos o design dos padrões e estrutura do teste a desenvolver no contexto desta investigação nacional. Apresentamos os modelos seleccionados. O Repositório Alberta, dentro da lógica do 'Alberta Education', apresenta um conjunto de padrões ('benchmarks') que funcionam como um sistema equilibrado de dimensões que abrangem vários níveis de proficiência e de ensino. É o sistema que temos vindo a abordar com maior enfoque dado o conjunto de padrões numa matriz alargada de descritores e exemplos sobre os níveis de proficiência por ano de escolaridade. Alguns destes descritores e referência a testes encontram-se também no modelo TOEFL (Test of English as a Foreign Language) que visa a aprendizagem do Inglês como LE (não como Língua Segunda) e disponibiliza testes internacionalmente reconhecidos para instituições educacionais que pretendam avaliar e certificar competências na área do Inglês como LE. O Brasil adotou os testes de TOEFL e aplica os procedimentos de forma a garantir a certificação e avaliação rigorosa dos brasileiros aprendentes de Inglês como LE. Por outro lado, o TESOL (Teachers of English to Speakers of Other Languages) é um dos modelos que prescreve princípios orientadores de avaliação de proficiência em Inglês como Língua Segunda. O Idea Proficiency Test (IPT) é um dos testes indicados por Alberta Education e que se baseia nos descritores do TESOL.

A bateria Minnesota Language Proficiency Assessments (MLPA) é um projeto do Center for Advanced Research on Language Acquisition (CARLA) e apresenta um conjunto de testes de avaliação da proficiência linguística em espanhol, alemão e francês, sendo que segue o modelo do American Council on Teaching of Foreign Languages (ACTFL). Apresenta, portanto um perfil distinto dos testes e baterias que vêm sendo referidas pois não aborda a aquisição do Inglês como LE ou L2. Os seus procedimentos e padrões de avaliação e aplicação de instrumentos seguem um perfil mais direccionado para a aquisição de línguas minoritárias, embora as competências medidas sejam sempre as quatro principais: escrita, oralidade, compreensão oral e escrita. Por sua vez, o Alaska Department of Education & Early Education disponibiliza um repositório de recursos (essencialmente fichas de identificação sociolinguística e documentos orientadores, para professores e pais) e informações na área de L2 mas referindo-se à população escolar imigrante como 'Limited English Proficient students'. Estes são submetidos a testes como o "WIDA-ACCESS Placement Test" (desenvolvido pelo Centro de Linguística Aplicada), dirigido para vários anos escolares (até ao 12.º ano) e um dos 'screening tests' mais referidos no contexto do modelo de inclusão praticado nas escolas do Alaska. Uma particularidade a destacar quanto a este modelo é a recomendação da sua aplicação no segundo semestre do ano letivo, além do apelo à preparação dos professores para a aplicação do instrumento.

O International Language Testing Association (ILTA) apresenta formas de medida com ênfase na linguística aplicada, na lógica da investigação fundamental mas com ligação aos profissionais em Educação. O ILTA assevera as instruções e normas na construção de testes de avaliação de competências linguísticas, sendo que seguiremos os procedimentos indicados, assim como conciliamos com os indicadores de ALBERTA sobre os procedimentos específicos das folhas de registo e dos padrões para cada nível de proficiência. No ILTA encontramos um importante corpus actualizado de trabalhos de investigadores especificamente na área da avaliação de linguagem, sendo que vários recursos apontam para investigações válidas na área das normas de construção de testes para medir competências várias em linguagem segunda. Os materiais disponibilizados, todavia, não são tão completos e variados quanto os recursos e documentos disponíveis na plataforma de Alberta Education.

Gascoigne & Parnell (2013) abordam uma análise muito pertinente sobre aspectos formais das tarefas de avaliação e aprendizagem em L2, mas nos contextos americano e europeu.

Constata-se que o Quadro Comum de Referência (Comissão Europeia, 2001) apresenta padrões para a avaliação e aprendizagem de línguas na Europa, ao passo que os autores americanos evidenciam outra panóplia de padrões no mesmo sentido, mas em escolas dos E.U.A. No entanto, é curioso como a literatura ainda apresenta lacunas no domínio europeu (e nas línguas europeias enquanto línguas alvo de aprendizagem no contexto americano) no que respeita à reflexão e revisão de testes na área de L2, sendo um problema a concretização das tarefas nos termos desejáveis,

Given the tremendous amount of professional literature devoted to tasks, either as a vehicle for focus on form or as a subject in its own right, it was surprising to find so little mention or use of tasks in the 18 commercially-available, post-secondary French materials reviewed. Indeed, only three out of the 18 post-secondary French textbooks reviewed and published in the U.S. between 1996 and 2012 (...). The tremendous variability between what we say about tasks and what we do in terms of the actual success of these applications—at least as represented in many post-secondary French textbook programs—is disappointing. One recommendation stemming from this admittedly anecdotal investigation is that, instead of taking the introductory descriptions of commercially-available materials at face value, language educators in the U.S. (in French or any other language) might be well served by borrowing the Common European Framework of Reference Guide for User's six-step auto-reflection tasks and four materials selection questions presented above. (...). By sharing materials, techniques, and guidance as evidenced in the Standards (in the U.S.) and the Frameworks (in Europe), we can move one step closer to our goals of multitasks, multiskills, and multiconnections.” (p. 98).

Com base nos descritores facultados por estes quadros e padrões é possível conceber níveis de dificuldade para as tarefas (os testes) e os itens dos questionários, o que já resolve parte da estrutura da avaliação pretendida, sendo que seguimos o procedimento sugerido por Alderson e Banergee (2002), onde são indicados alguns dos quadros teóricos referidos anteriormente,

views of difficulty in reading in a foreign language could be explored by looking at how test developers have specified their tests at different levels of proficiency, and refers to the Cambridge suite of tests, the Council of Europe Framework, and the frameworks used in national assessments of foreign language ability. (p. 84)

Além dos padrões e descritores que as entidades de avaliação (exemplo: ALBERTA, ILTA) providenciam à comunidade académica e não académica, os testes específicos de avaliação de competências e proficiências também incorporam indicadores que podemos aproveitar para orientar novas folhas de registo e de pontuação de dados. Em Portugal, os documentos orientadores e testes de avaliação diagnóstica nesta área apresentam tipos de registo de dados a considerar sobretudo na tarefa designada como “reconto” em que é avaliada a ordem dos eventos da narrativa, a utilização de vocabulário adequado, a construção dos enunciados frásicos, a previsão de ‘situações sincopadas’, de ‘não realização da atividade’ e de indicação de sinónimos em vez dos nomes pretendidos para a nomeação das imagens; e, ainda, a determinação de três graus de complexidade nas fichas de registo (DGE, 2009). No entanto, o tipo de pontuação e de estrutura dos testes apresenta várias limitações. A avaliação do vocabulário e as respetivas listas de palavras não seguem os critérios de frequência lexical ou, pelo menos, não há referência na explicação da construção dos específicos exercícios. A determinação da pontuação (scores) não se encontra justificadas em termos psicométricos, mas sobretudo apresenta uma lacuna a evidenciar: na avaliação global os níveis de proficiência são associados a intervalos de escalas que são iguais para todos os anos de escolaridade, além do facto dos valores serem também equivalentes para os domínios da proficiência oral e escrita, em todos os testes desenvolvidos e aplicados. Não se percebe se os docentes, que administraram os testes, entenderam a relação entre os valores estipulados pelos autores (DGIDC 2012) e os níveis de proficiência determinados pelo Quadro Europeu Comum de Referência para as Línguas (Comissão Europeia, 2001). Não houve validação nessa correspondência ou algum racional descrito para a explicitar, apesar das provas elaborados pelo ILTEC (encomendadas pela DGIDC) se apresentarem como uma “proposta inicial” (DGIDC, 2012, p. 3). Também não se clarifica propriamente se os professores entendem quando se referem a “posicionamento rigoroso” nos níveis de proficiência, pois os scores não foram informados quanto ao seu racional e fundamentação teórica, ainda que sendo provas criadas originalmente. Quanto à opinião sobre os instrumentos de registo completos,

disponibilizados pela DGIDC (2012), também se concorda com este facto na medida em que são fichas de registo muito bem organizadas e atentas a situações da fonética que os locutores de outras línguas frequentemente apresentam de forma muito variada e particular.

Os testes diagnósticos disponíveis em Portugal, na área de PLNM, apresentam um formato próximo (sobretudo nas imagens e nos temas) da estrutura das provas do ACCESS for ELLS (2008). Os estímulos são essencialmente visuais (exemplo da soletração, nomeação oral e escrita, produção escrita), o que é também frequente neste tipo de provas. Um dos aspetos positivos reside numa boa referência à avaliação do acento no discurso fonético (perceção do desconhecimento dos alunos face às letras do alfabeto e aos contextos fonéticos de nasalização) e à avaliação de coesão e pragmática (na relação com os géneros textuais). Limitações voltam a encontrar-se no tipo de textos escolhidos para os diferentes anos de escolaridade uma vez que são muito similares, assim como os estímulos visuais referentes a imagens. Aliás as listas de palavras, dispostas nas provas dos diferentes níveis de ensino, são também muito similares, portanto, com pouca diferenciação lexical e semântica. Há diferenciação dos anos de escolaridade mais baixos em relação aos mais avançados como o 3.º ciclo e ensino secundário, mas entre estes níveis não há distinção na estrutura de avaliação. Nos indicadores dos ‘benchmarks levels’ do Alberta Education (2010) os níveis de escolaridade são todos distinguidos por nível e de forma intensiva e detalhada. Desta forma serão considerados muitos dos seus indicadores em provas que se encontram a ser construídas no âmbito deste projeto de pós-doutoramento. Para informação do professor, os testes em causa não distinguem as competências por áreas de comunicação (Alberta Education, 2010, p. 5). A aplicação dos testes decorreu durante dois meses o que pode ser um período (método transversal) reduzido dada o volume de escolas (163 solicitaram) que aderiram à aplicação dos testes apesar de nem todas terem efetivado a aplicação e o preenchimento do questionário que foi considerado para a ‘validação’ das provas e deste projeto da DGIDC. Esse questionário é a base do estudo de validação conduzido pelo ILTEC e DGIDC (DGIDC, 2012) mas constituiu-se como um inquérito à perceção dos docentes face à dificuldade e vantagens da aplicação dos testes em causa. Os testes não foram validados em termos de desempenho o que não permite calibrar os scores, nas provas individuais e sobretudo na classificação global para cada ano de escolaridade. Por outro lado a análise das percepções dos docentes é maioritariamente de ordem qualitativa e incide sobre aspetos de compreensão dos enunciados e das instruções (e do tempo de aplicação) mas na perspectiva dos professores e não dos alunos visados. No entanto, com base nessas observações a complexidade de enunciados e outros aspetos foram aperfeiçoados nas provas, conforme informado pela DGIDC (2012). De forma curiosa, os resultados desse questionário permitiram também perceber que as escolas continuam sem utilização significativa dos documentos da DGE, o que Nunes (2010) também assevera na sua investigação de mestrado. Neste estudo a autora recorreu aos testes diagnósticos de 2005 da DGIDC (Pascoal & Oliveira, 2005) na medida em que eram os que se encontravam disponíveis, na mesma plataforma da DGIDC. No entanto este estudo apresenta um perfil de intervenção e não propriamente de avaliação diagnóstica pois indicadores de validação das provas também não se referem, sendo que a intervenção e a aprendizagem são os objectivos que mais se destacam na leitura do estudo. A construção de instrumentos, em complementaridade com os projectos existentes (DGIDC), é uma demanda que se mantém em nível de urgência.

Concentrando-nos no último ponto, no domínio da evidência científica, alguns aspetos teóricos serão considerados. Especificamente no contexto teórico da “dupla via” importa-nos abrir aqui uma reflexão sobre a alternativa da chamada ‘leitura por analogia’ (Gillon) que poderemos entender como uma leitura e aquisição de conhecimento vocabular de forma económica e com menor esforço, atendendo a uma espécie de reciclagem de tipos de léxico que permite reconhecer palavras que têm os mesmos padrões fonéticos. Aplicaremos este princípio teórico no contexto de aquisição e desenvolvimento de linguagem segunda (todo o tipo de língua ou línguas que o sujeito adquire com proficiência, não se refere apenas a um idioma). A criança que já desenvolveu a consciência do nível intrassilábico pode mais facilmente deduzir léxico novo ou pseudoléxico a partir do que já conhece na medida em que os seus padrões fonológicos se encontram suficientemente amadurecidos para adquirir léxico por analogia (Goswami & Mead, s.d.). Atendendo aos objectivos do trabalho de investigação em curso – desenvolvimento de uma bateria de avaliação diagnóstica compreensiva – centramo-nos na teoria de que um sistema de relações ao nível semântico, ortográfico e, obviamente,

fonológico: “1. Knowledge of language consists of generalizations made over linguistic pattern complexes (LPCs) (...). 2. Language acquisition is an auto-associative process (...) and on this basis associations are built up between the microfeatures of LPCs (via the hidden units)” (Gasser, 1990, p. 12). Todos esses níveis e processos associativos implicam, então, que os exercícios de avaliação devam abranger os mesmos níveis e explorar as competências dos locutores com o objetivo de determinar classificação de proficiência e mesmo de casos de dominância linguística.

Por outro lado, na interpretação das respostas, na perspectiva do investigador/autor do teste, e na visão dos professores que utilizarão os instrumentos, há particularidades no processo (cognitivo) de aquisição da Língua Materna (LM ou L1) e da Língua Segunda (L2):

There are three ways in which second language acquisition may differ from first. (...) 1. L1 patterns may transfer to L2 (and vice versa). 2. Neurophysiological changes or cognitive developments not related specifically to language may limit the learner's ability to acquire language or may predispose the learner to particular acquisition strategies. 3. Contextual factors, such as the acquisition setting or the communicative demands placed on the learner, may affect acquisition (Gasser, 1990, p.13).

A estruturação dos exercícios que compõem o teste diagnóstico segue preditores constituídos pelas premissas das investigações internacionais anteriores. Por exemplo a nossa preocupação em termos um exercício sobre inteligibilidade e compreensibilidade do discurso (com recurso a ficheiros de som com e sem sotaque, em Português) prende-se com a premissa de que a génese de um léxico materno é baseada na avaliação do input ouvido de forma isolada e, se se afirma que a percepção fonética (e fonológica) determina completamente a organização lexical, a dificuldade com novas assimilações lexicais (em Língua Segunda) poderá ser determinada por essa percepção. Contudo, neste caso sugerido, os aprendentes mais novos apresentariam menos perturbação a este nível na medida em que a sua percepção e produção fonéticas são consideradas mais fiéis ao código alvo. A percepção fonética, por outro lado, enquanto ferramenta de compreensão e aprendizagem da pode ser bloqueada com os efeitos competitivos que são mais ou menos exacerbados de acordo com a similaridade dos códigos em confronto. Outro aspecto é a tendência de assimilação dos sons estranhos a identidades do sistema fonológico e fonético maternos, sobretudo no que diz respeito aos adultos. Nem sempre dois códigos similares poderão auxiliar no processo de aprendizagem das duas fonologias na medida em que o efeito competitivo é mais evidente e, portanto, prevê-se mais prejudicial ao processo de descodificação. Por outro lado, temos ainda o efeito ortográfico, em que se verifica a influência do conhecimento abstracto e episódico no processamento lexical, ou seja, não existe apenas activação estimulada pelo input sonoro. As representações fonológicas serão possivelmente determinadas pelo efeito fonético e ortográfico. Exercícios com estímulo ortográfico são os mais vistos no teste que se encontra em aplicação. O acesso mais ou menos predominante no que respeita às duas (lexical e fonológica) vias anteriormente enunciadas depende do contexto linguístico a que se aplicam. O procedimento fonológico ou lexical varia em grau de importância de acordo com o tipo de Língua em causa, pois dependerá de qual tipo é instruído, assim terá repercussões nos ritmos de desenvolvimento da consciência fonológica e, conseqüentemente, das competências de leitura e escrita (Figueiredo, 2010).

No estudo anterior (Figueiredo, 2010), uma das limitações consideradas foi a necessidade de maior replicação de testes em amostras mais homogêneas, em termos de nacionalidades e línguas maternas dos locutores, assim como o questionário inicial dever conter mais informação sobre a caracterização da amostra: o nível socioeconómico dos pais/tutores e se o locutor é dextro ou esquerdino. Estas especificidades poderão determinar interpretações fundamentais tais como sobre o processamento cognitivo na apreensão de sons a partir da exibição dos testes com ficheiros sonoros. Quanto ao critério de homogeneidade da amostra, naturalmente populações tão caracterizadas pela diversidade como o caso das migrantes torna difícil restringir a variável nacionalidade ou língua materna. Mas, neste estudo excluímos todos os locutores oriundos de PALOP's de forma a evitar avaliar, no mesmo contexto, alunos que falam crioulos de base lexical portuguesa e alunos que falam outro tipo de línguas. Cingimo-nos essencialmente a locutores de línguas eslavas e línguas indo-arianas, também atendendo à sua representatividade na escola portuguesa. Replicamos no nosso teste alguns exercícios já

utilizados na aplicação do instrumento de investigação de doutoramento, como o caso do exercício de identificação silábica, intrassilábica e fonémica. Os resultados anteriores (Figueiredo, 2010) revelaram que não há, de facto, assimetria entre os grupos, de acordo com os níveis silábico, intrassilábico e fonémico. Interessante observar que no grupo de sujeitos nativos (portugueses) houve alguma linearidade, por exemplo, a criança tem melhor performance no teste de divisão silábica e mesmo no de identificação de rima, em relação ao grupo adulto. No grupo experimental (com alunos com experiência migratória) verificamos que os sujeitos adultos dominaram no âmbito do nível da sílaba, contudo não da rima. Há, assim, uma competência a desenvolver-se (considerando o adulto que está na fase de iniciação da sua aprendizagem fonológica em L2) e que não significa que necessariamente siga a ordem típica na aprendizagem de língua materna, contudo também não significa por isso que haja correlação inversa proporcional. O que poderemos sugerir é que não há ordem necessária a ser respeitada, os níveis vão sendo adquiridos de forma até possivelmente aleatória e poderá ser que a intervenção da estratégia de ensino influencie esse tipo de ordenação, na medida em que notamos alguma flexibilidade nos termos em que a aprendizagem se processa. Este é um tipo de hipótese que colocaremos de novo em escrutínio.

## REFERÊNCIAS

- Alberta Education, "Assessment tools & strategies, language proficiency assessment": <http://www.learnalberta.ca/content/eslapb/languageproficiencyassessment.html> (2012).
- Alberta Education, "English as a second language": <http://education.alberta.ca/teachers/program/esl.aspx> (2010).
- Alderson, J., & Banerjee, J. (2002). Language testing and assessment (Part 2). *Language Teaching*, 35, 79-113.
- Conselho da Europa. (2001). *Quadro Europeu Comum de Referência para as Línguas: Aprendizagem, ensino, avaliação*. Porto: Edições Asa.
- Direção-Geral da Educação (2012). "Teste de Diagnóstico de Português Língua Não Materna" (1.º ano – 12.º anos), <http://www.dgdc.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.
- Gasser, M. (1990). Connectionism and Universals of Second Language Acquisition. *Studies in Second Language Acquisition*, 12, 179-199.
- Gillon, Gail T. (2004). *Phonological Awareness: From Research to Practice*. NY: Guilford Press.
- Højsholt-Poulsen, L. (s/d). *Repositórios Educativos: Em Busca dos Sucessos Europeus*. UNI.C, Centro para a Educação e a Investigação em Tecnologias de Informação, Ministério da Educação, Observatório dos Recursos Educativos, <http://www.ore.org.pt/observatorio/estudos.aspx?contentid=AE7B1FDFF8EC0C4DE0400A0AB80033F8&lang=pt>.