

# The Language Instruction Scheme for Graduates with a Withheld Degree

**Zoltán Lukácsi**

*Euroexam International, Budapest, Hungary*

*Budapest Business School, Budapest, Hungary*

*lukacsi74@gmail.com*

## 1. Introduction

University undergraduates in Hungary are required by law to have at least one state-accredited foreign language examination certificate at level B2 if they are to gain their degrees (Ministry of Human Capacities Decree 18.). Over the past decade and due to excessive standardization, language instruction at the tertiary level has undergone unfavourable curricular changes. A highly contentious issue is the heavy reduction in the time spent on teaching and learning languages. Degree courses at the Budapest Business School schedule 156 contact lessons, whereas Corvinus University offers 104, and some vocational further education programs at the same institutions only provide 91. As a rule, the cost of language instruction is covered by the tuition fee. However, as the student is held responsible for meeting the academic requirements, most universities are reluctant to invest into language teaching and offer language classes as extra-curricular activities that incur additional expenses.

Input features high on the comprehensive lists of factors that influence L2 learning and the rate of acquisition (Ellis, 2004; Littlewood, 2004). One, albeit crude, way of quantifying input is to express it in terms of guided learning hours, i.e. tutor-led contact lessons. Even though de Jong (2009) harshly criticized the Cambridge English Language Assessment webpage (<http://www.cambridgeesol.org/exams/exams-info/cefr.html>) for the overly simplistic view of L2 learning pace, it still remains the only comparative measure to guide curriculum design. However basic the recommended 150-200 hours might be to reach level B2 from B1, even the most robust instruction program in Hungarian higher education falls behind these recommendations.

Besides, without an entrance examination or a filter test, universities have no information available about students' L2 level at the outset of teaching. Previous research conducted locally at the Budapest Business School (Lukácsi, 2015, 2016) shows that the curriculum cannot adequately cater for the L2 needs of the majority of university undergraduates, and roughly a third of the freshmen will be unable to meet the English language certificate requirement. Indeed, on the country level in 2014, some 50,000 individuals could not collect their degrees through lack of a language exam (Soós, 2015).

## 2. The language instruction scheme

In 2014, the Ministry of Human Capacities commissioned the National Employment Coordination Agency (OFA NKft.) to develop, organize, and control a language

instruction scheme for university graduates with a withheld degree. Private language schools were invited to apply for funding from the 10,000,000€ budget to run free classes in English (72%), German (25%), and French (2%). The intensive courses were arranged into 120, 180, 240, and 360 hour blocks with 6 to 20 contact lessons a week (“Így jelentkezhettek,” 2015). Applicants were expected to (a) have completed their university education, (b) have successfully passed their final examination formerly known as the state examination, and (c) be at level B1 in the language in question. Under the terms of their contract, the language learners agreed to pass a B2 exam by the end of 2016 in return for the free instruction.

Thus far, despite the volume of the venture, no scientifically sound evaluation of the language instruction scheme has been conducted. Publicists and government officials depict the project as a total flop with a pass rate of 15% (G. Tóth, 2015) or as a triumphant success with that of 90% (Soós, 2015; Tarnai, 2015). These articles and internal reports are in stark contrast to viable psychometric expectations and current state language examination accreditation requirements alike. Dávid (2014) claims that a realistic pass rate in a language exam should fall between 35 and 65 per cent. The Accreditation Manual (2016) rules that item difficulty must be between 30 and 70 per cent, which implies that a test paper will also lie within these extremes.

### **3. The study**

#### **3.1. Aims and research questions**

The study reported here aimed to answer three research questions.

RQ 01: What is the pass rate among candidates with a withheld degree?

RQ 02: Do candidates with a withheld degree produce the same pass rate as the rest of the candidature?

RQ 03: How to explain potentially different degrees of success between candidates with a withheld degree and the rest of the candidature?

#### **3.2. Method**

The validity of such a comparative analysis hinges on the existence of two distinct groups of test takers: (a) regular candidates and (b) candidates with a withheld degree. In the rest of the paper, the term “regular candidate” will be used to describe test takers who were not enrolled in the language instruction scheme. As a preliminary requirement to any meaningful comparison, live examination data were first screened to see if a longitudinal analysis was even possible between pre-2015 and 2015 candidates. Therefore, evidence needed to be collected to support the claim that candidates with a withheld degree were altogether new to the examination system. Given that they had to indicate in a checkbox on the language exam application form if they had enrolled in the language instruction scheme, a simple numerical comparison of annual candidature together with identity checks were sufficient. As a result, the validity of the claim that candidates with a withheld degree were different from regular Euro candidates was established.

### 3.3. Participants

The participants were 12,315 candidates at level B2 who sat for the Euro examination in 2015. The vast majority took the monolingual examination, i.e. tests in listening, reading, speaking, and writing (N = 11,748) without the mediation paper. The sample consisted of two subsamples: (a) candidates with a withheld degree (N = 1,920) and (b) regular examinees (N = 10,395).

### 3.4. Design and instruments

Euroexam International is one of the few certified L2 centres in Hungary that use a data collection design which enables direct comparison of test and examinee statistics longitudinally. In a non-equivalent groups anchor test (NEAT) design (Kolen, 2007), each test administration contains a number of repeated items with known characteristics and, further, a stable ability standard is applied to ensure that the same amount of the latent trait is required in consecutive test periods. In essence, it means that regardless of when the candidate sits for the exam and which tasks they have to complete, their results will be fair and directly comparable. For the estimation of item difficulty and language ability, the item responses were processed in the One-parameter Logistic Model: OPLM (Verhelst, Glas, & Verstralen, 1995). To answer the research question about differential degrees of success, Profile-g (Verhelst, 2012) was used to model the statistical profiles of the two cohorts.

### 3.5. Results and discussion

In terms of the results, reported score means are compared first in the four test parts: listening, reading, speaking, and writing. Next, the pass rates in each part will be presented. Since the Hungarian system of accreditation uses a mixture of the conjunctive and the compensatory approaches (Kaftandjieva, 2004) when combining the test papers into the final examination result, overall achievement and success will be reported separately.

### 3.6. Percentage scores and pass rates

Test paper means are calculated as simple arithmetic averages of the reported scores in the two groups of test takers (Table 1). Officially, these are expressed as truncated percentage values even though the scores do not constitute a ratio scale, and the extremes do not mean total lack or possession of the latent trait in question.

Table 1 Mean Test Paper Scores of Regular and Withheld Degree Candidates

| Candidates | Listening | Reading | Speaking | Writing |
|------------|-----------|---------|----------|---------|
| Regular    | 60.5097   | 62.2977 | 64.9267  | 59.4248 |
| Withheld   | 51.6312   | 60.2512 | 56.5986  | 55.3920 |
| Grand mean | 59.0627   | 61.9612 | 63.5696  | 58.7617 |

Despite the methodological discrepancies discussed previously, Table 1 demonstrates the difference in attainment between regular test takers and those with a withheld degree. In 2015, candidates with a withheld degree fell behind the grand mean, but particularly the regular examinees in every respect. Listening results showed the biggest difference, followed by speaking, writing, and then reading. A viable interpretation of these figures is to claim that the language instruction scheme was most efficient in teaching language learners to read.

An analysis of the pass rate yields a more criterion-referenced overview of the examination results. In this comparison, the groups are subdivided into two categories according to success on the test paper (Table 2). The pass rate then is the percentage of successful candidates in each group.

Table 2 Pass Rate (%) of Regular and Withheld Degree Candidates

| Candidates | Listening | Reading | Speaking | Writing |
|------------|-----------|---------|----------|---------|
| Regular    | 54.9312   | 56.8391 | 67.6763  | 49.4340 |
| Withheld   | 36.8171   | 51.6667 | 46.7933  | 37.5269 |
| Grand mean | 51.9791   | 55.9887 | 64.2733  | 47.4764 |

The test paper means in Table 1 already described them as underachievers, but the pass rates in Table 2 depict candidates with a withheld degree as substantially weaker than regular candidates. With the exception of reading, they were more likely to fail than to pass, which was particularly prominent in listening and writing. The reason for the increasing difference between the two cohorts resides in the fact that the consecutive administrations are only equated at the standard. The stable pass mark is at 60 points on the reporting scale of 0 to 100. However, no correct responses are always worth 0 and completely flawless responses are always 100 irrespective of the difficulty of the set of tasks. While the same amount of ability is required for a pass from each candidate regardless of exam period, it also implies that nowhere else on the reporting scale will two similar numbers signify the same. As a result, arithmetic averages will distort genuine differences in ability to varying degrees.

As Government Decree 137 / 2008 2.§ (4) b) rules that in terms of type a language exam can be (a) oral, (b) written, or (c) complex if (a) and (b) are combined, the most authoritative criterion when evaluating the language instruction scheme is the overall success (Table 3).

Table 3 Overall Means and Pass Rates of Regular and Withheld Degree Candidates

| Candidates | Mean    | Pass rate (%) |
|------------|---------|---------------|
| Regular    | 60.7989 | 52.9248       |
| Withheld   | 56.0775 | 38.2813       |
| Grand mean | 60.0627 | 50.6415       |

As Table 3 shows, collating the test paper results into an overall score distorts and misrepresents individual differences resulting in a reduced distance between the two groups in the mean. By contrast, the pass rate is robust enough to such a distortion and reveals that they were roughly 15% less likely to be successful than

regular candidates. Grounded in empirical evidence, the response to RQ 01 is that 38% of learners from the language instruction scheme obtained a certificate in English from Euroexam International in 2015. This is significantly lower than the 53% pass rate observed among regular candidates.

### **3.7. Reasons for the different degrees of success**

In the next part of the study, the statistical profiles of the two contrasted cohorts on the objectively scored test papers will be presented. Profile analysis is similar to differential item functioning (DIF) analysis in that it can detect systematic deviations from model predictions (Verhelst, 2012). However, while DIF works on the item level, profile analysis partitions items into categories, e.g. tasks, and creates deviation profiles accordingly. A deviation profile shows the difference between the expected profile and the observed profile. As such, all deviation profiles will necessarily sum to zero.

Instead of reporting on the results from each administration separately, the findings will be presented from an aggregate where all the candidates from the same group are combined essentially reducing the method to a pairwise comparison. Of particular importance are the deviation profiles of candidates with a withheld degree to see how balanced their performance was and, eventually, how to explain their poor pass rate. In essence, such an analysis can reveal if a group performed unexpectedly well (or poorly) on a task type when compared to what would be predicted based on the total score on the test paper.

Following the Detailed Specifications (Euro Examinations, 2009) as a blueprint, the listening paper contained three task types: (a) short conversations, (b) making notes, and (c) radio programme. A thorough discussion of what each item in the three test tasks aims to measure is impossible in a publication such as this one, but the general principle is to proceed from a global comprehension to understanding details including stance and attitude, and from short, simple passages towards longer, more complex texts. None of the tasks tapping receptive skills require production of language; therefore, guessing could always play a part, especially so in the radio programme, a three-option multiple choice task.

Test paper reliability was always sufficiently high ( $\alpha \geq .75$ ), and DIF was not present. The data were processed in the OPLM-model with the geometric mean of discrimination indices set to  $g = 3$ . Model fit was always checked and deemed acceptable for practical purposes. Mean person parameter estimates quite bluntly depicted candidates from the language instruction scheme as less able. Figure 1 presents the statistical profiles of withheld degree test takers and regular candidates broken down to task type.

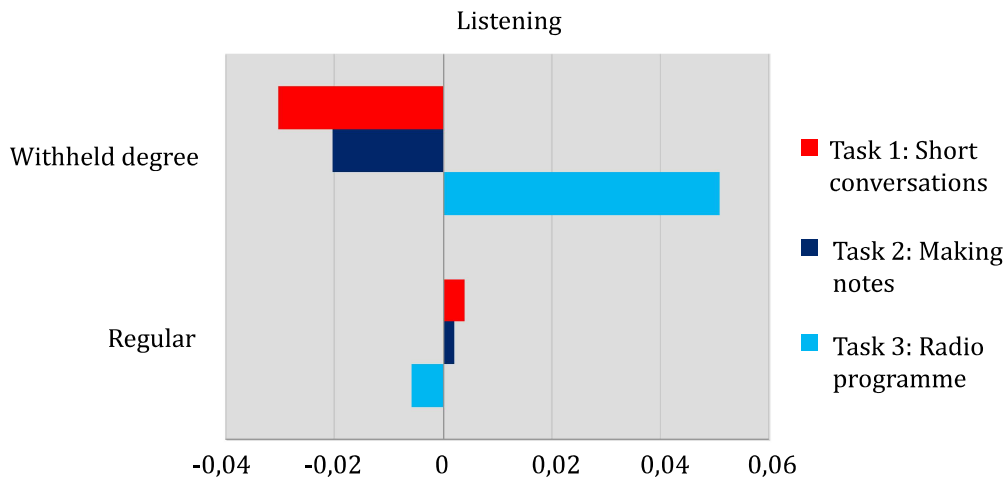


Figure 1. Statistical profiles on the listening paper

In Figure 1, the bars associated with regular examinees are much shorter than in the other group primarily because the sample sizes were not equal. The probability of scoring higher than expected is particularly conspicuous in the withheld degree group on Task 3. On the other hand, the same candidates were more likely to score low on Task 1, a task focusing on more detailed, global comprehension. As model fit was acceptable, the variance in the scores was attributed to differences in the latent trait rather than guessing, test-wiseness or other sources of noise. Arguably, the language instruction scheme successfully prepared these candidates for finding specific, predictable information without understanding longer stretches of spoken text. The two task types where withheld degree test takers underperformed both required comprehension and processing of longer passages and, as such, also placed a heavier burden on working memory capacity.

The reading paper also contained three task types: (a) paragraph headings, (b) scanning, and (c) multiple choice reading. The first two tasks were matching, whereas the last one was a four-option multiple choice task. The psychometric properties of this test paper were similar to the listening part. Figure 2 displays the reading profiles of withheld degree test takers and regular candidates.

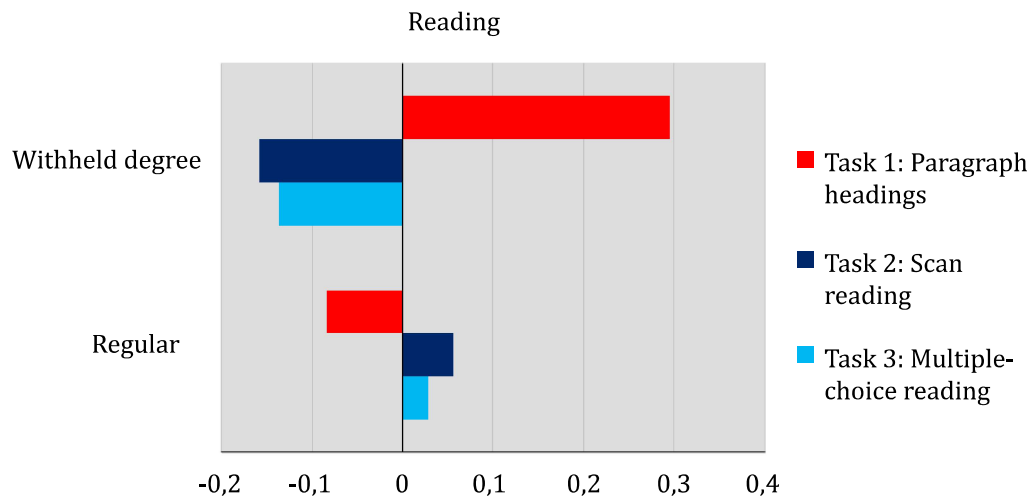


Figure 2. Statistical profiles on the reading paper.

As Figure 2 shows, similarly to the listening profiles, unexpected success on the reading paper was also largely attributed to a single task. This time Task 1: Paragraph headings yielded much better scores than anticipated on the basis of total scores. In this task type, the candidate is expected to choose a heading from a given pool of eight options to six short paragraphs of about 70 words each. The language instruction scheme prepared these language learners for the surface processing of short printed texts. Scan reading and the multiple choice task type both require processing longer passages, which goes beyond local comprehension. It is also important to point out that paradoxically Task 2 is not scanning as the name would suggest but rather search reading (Weir & Khalifa, 2008). While both are classified as expeditious reading, scanning is selective reading at the word level for specific items, whereas search reading never aims for exact word matches.

#### **4. Conclusion and implications**

Contrary to the enormous stakes of foreign language examinations in Hungary, the role of L2 teaching and learning in higher education is secondary at best. A number of institutions delegate the responsibility of mastering an L2 to the student and for financial reasons refrain from providing L2 education as part of the curriculum. An unwelcome result of this educational policy is reflected in the gradually increasing number of withheld university degrees. By 2014, the number of people unable to collect their degrees had reached 50,000. In response to the growing need for L2 instruction, the Ministry of Human Capacities launched the language instruction scheme providing ample financial support for private institutions to run intensive exam preparation courses. Despite the volume of the budget, no scientifically sound investigation has been launched into evaluating the project.

In this comparative analysis, live examination data from 2015 were reviewed. Even though the sampling applied was comprehensive in the sense that it included the entire population of test takers in Euroexam International, the results only generalize to other certified language centres as far as the efficiency of the language instruction scheme and the level standards in the various language testing systems can be conceived of as constant. With these contentions, the study found that the language instruction scheme helped 38.2813% of the graduates with a withheld degree to their L2 certificates and, more importantly, to their university qualifications.

Any interpretation of such a low pass rate will require further judgement. The language instruction scheme was undoubtedly a success story if the position taken is that each additional degree is a positive result. However, considering the vast amount of resources, the number and intensity of the classes especially in relation to university L2 instruction, and the quantity of the remaining withheld degrees, the language instruction scheme failed to reach its major goal. The language school courses were successful in cramming learners for exam tasks where local comprehension was the focus. The study also found that withheld degree candidates found it hard to deal with longer passages, particularly when global understanding of deep text structures was targeted.

The findings from this study provide empirical evidence in support of previously published recommendations. Nikolov (2011) pointed out that universities should systematically develop students' L2 and suggested that content and language integrated learning should be introduced. Studying a subject per semester in an L2

would result in increasing levels of motivation. Further, such authentic language use would help the language learner acquire expert knowledge and skills eventually leading to learner autonomy. Unless educational policymakers invite professionals in curriculum development, syllabus design and language planning to assist in redesigning the current system of L2 instruction, universities will continue to produce graduates with withheld degrees.

## 5. References

- Akkreditációs kézikönyv* [Accreditation manual]. (2016). OH: NYAK. Retrieved November 24, 2016, from <http://www.nyak.hu/nyat/doc/ak2016/ak2016.htm>
- Dávid, G. (2014). Mitől lesz jó egy nyelvvizsga? [What makes a good language exam?]. *Iskolakultúra*, 24(4), 3-25.
- De Jong, J. (2009, June). *Unwarranted claims about CEF alignment of some international English language tests*. Paper presented at the meeting of EALTA, Turku, Finland.
- Ellis, R. (2004). Individual differences in second language learning. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 515-552). Oxford: Blackwell.
- Euro Examinations. (2009). Euro detailed specifications. Budapest: EuroExaminations.
- G. Tóth, I. (2015). Nyelvi káosz: Botrányba fulladt a diplomamentő akció [Language chaos: The language instruction scheme ends in scandal]. *HVG*. Retrieved November 28, 2015, from [http://hvg.hu/itthon/201548\\_botranynba\\_fulladt\\_diplomamento\\_akcio\\_nyelvi](http://hvg.hu/itthon/201548_botranynba_fulladt_diplomamento_akcio_nyelvi)
- Government Decree 137. (2008). Retrieved December 11, 2015 from [http://www.nyak.hu/doc/jogszab/137-2008\\_EN.pdf](http://www.nyak.hu/doc/jogszab/137-2008_EN.pdf)
- Így jelentkezhettek az ingyenes nyelvtanfolyamra [This is how you can apply for admission to the free language courses]. (2015, July). *EduLine*. Retrieved July 31, 2015, from [http://eduline.hu/nyelvtanulas/2015/7/31/diplomamento\\_kerdesek\\_XJ50NM](http://eduline.hu/nyelvtanulas/2015/7/31/diplomamento_kerdesek_XJ50NM)
- Kaftandjieva, F. (2004). Section B: Standard Setting. *Reference Supplement to the Preliminary Pilot Version of the Manual for Relating Language Examinations to the CEFR for Languages: Learning, Teaching, Assessment*. Strasbourg: Language Policy Division.
- Kolen, M. J. (2007). Data collection designs and linking procedures. In N. J. Dorans, M. Pommerich, & P. W. Holland (Eds.), *Linking and aligning scores and scales* (pp. 31-55). New York, NY: Springer.
- Littlewood, W. (2004). Second language learning. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 501-525). Oxford: Blackwell.
- Lukácsi, Z. (2015). *How a CEFR-linked test of ESP can inform educational policy*. Paper presented at the meeting of EALTA, Copenhagen, Denmark.
- Lukácsi, Z. (2016). Értékteremtés a turizmus-vendéglátás szakos hallgatók szaknyelvi képzésében – avagy: hogyan mérjük az oktatás segítésére [Added value in the foreign language instruction program for tourism and catering majors – or: how to test to assist education]. *Nyelvvilág*, 17, 5-18.
- Ministry of Human Capacities Decree 18. (2016). Retrieved December 11, 2016 from [http://net.jogtar.hu/jr/gen/hjegy\\_doc.cgi?docid=A1600018.EMM&timeshift=ffffff4&txreferer=00000001.TXT](http://net.jogtar.hu/jr/gen/hjegy_doc.cgi?docid=A1600018.EMM&timeshift=ffffff4&txreferer=00000001.TXT)
- Nikolov, M. (2011). Az idegen nyelvek tanulása és a nyelvtudás. *Magyar Tudomány*,



172(9), 1048-1057.

- Soós, A. (2015, November). *A diplomamentő programról* [About the language instruction scheme for graduates with a withheld degree]. Paper presented at the meeting of Certified Foreign Language Examination Centres, Budapest, Hungary.
- Tarnai, R. (2015, December). Sikeres volt a diplomamentő program [The language instruction scheme was successful]. *Pest Megyei Kormányhivatal*. Retrieved December 11, 2015, from <http://www.kormanyhivatal.hu/hu/pest/hirek/siker-es-volt-a-diplomamento-program>.
- Verhelst, N. D. (2012). *Profile-g: A computer program for profile analysis*. Tiel: Eurometrics.
- Verhelst, N. D., Glas, C. A. W., & Verstralen, H. H. F. M. (1995). *One parameter logistic model: OPLM*. Arnhem: Cito.
- Weir, C., & Khalifa, H. (2008). A cognitive processing approach towards defining reading comprehension. *Research Notes*, 31, 2-10.