ISSN: 2669-2325 (Online)
International Journal of Childhood Education • Vol. 2 • Issue 2 Open Access

Second Language Testing: A Psycholinguistic Approach

Sandra Figueiredo

Associate Professor of Psychology and Ph.D. Integrated Researcher of Psychology Research Centre (CIP), Autonomous University of Lisbon, UAL., Portugal
* **Corresponding Author:** sfigueiredo@autonoma.pt

Citation: Figueiredo, S. (2021). Second Language Testing: A Psycholinguistic Approach. *International Journal of Childhood Education, 2* (2), 26-31. https://doi.org/10.33422/ijce.v2i2.15

ABSTRACT
Cut-off analysis is very important in the Second Language testing domain considering the type of tools administered at schools and beyond. Verbal reasoning and lexical decision are crucial to be examined along with additional linguistic tasks in immigrant school population. 108 Portuguese immigrants, with ages between 7-17, from 20 countries and speakers of 21 mother tongues completed a laterality task (dichotic listening test) in order to understand if this type of task should integrate a full diagnostic test for non-native students. Participants answered the test in a single session lasting approximately 60 minutes (along with more tasks in a battery), in the same space that had been previously stipulated by the schools and on appropriately determined schedules. Dichotic Listening test assesses both cognitive and linguistic skills mainly focusing lexicon retrieval and semantic comprehension. Phonological dimension is also evaluated by the binaural detection test. Portuguese immigrants with an age of onset (age of acquisition of Portuguese as Second Language) under eight years were considered for this sample. The Receiver Operating Characteristic (ROC) analysis was carried out to define the cut-off points of different tasks. In the present study we will present the specificity, sensitivity and AUC for the dichotic listening. The score calculation took into account the laterality (right or left-handed) and also controlling for gender and age variables. The results pointed out to the predictive value of the dichotic listening task for right-handed, meaning that is able to distinguish proficiency levels and immigrants at risk (right-handed) concerning the evaluation and the learning process. These data highlight the discriminatory power of specific tests with specific sensitivity and specificity properties that could inform which areas are fragile attending the inter-individual variation.
keywords: dichotic listening task, psycholinguistics, second language, language deficits, assessment

1. Introduction

Scientific studies on evaluation of immigrant skills have been focusing, in Second Language (L2) testing area and attending to the lateralization in young individuals, the verbal lexicon, the fluency, the reaction time (Namaziandost et al. 2019; Suzuki & Kormos, 2019; Ulbrich & Wiese, 2018) and the dichotic listening (Fatahi et al., 2018; Hiraoka, 2018) in order to compare different age groups of immigrants. For that evaluation the cut-off points are important to be examined and established. Despite the lack of evidence available, specific tests such as the dichotic listening and respective cut-off scores should be validated for immigrant populations, considering each hosting country (Kuiken & Vedder, 2017; Hille & Cho, 2020) but also the ethnic groups (so, considering the country of origin and culture/learning/prior knowledge). The evaluation tests and reference scores will be understood only attending to the variables such as the case of ethnicity. The ethnic variable is of upmost importance in the cut-off analysis as we can observe for health or psychological assessment: different populations demand different tests (Saxena, Ambler & Cole, 2004; Wang & Wang, 2002).

Cut scores are currently determined by descriptive and regressive analyses or limited to "item response theory" analyses. Considering inter-individual factor, the statistical analyses should integrate a comprehensive data to understand the sensitivity and specificity properties of the

© The Author(s). 2022 Open Access. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author(s) and source are credited.

tests. Attending to this principle, we conducted our research with the receiver operating characteristics (ROC) analysis. By using this method, the cut scores are completed with models and maximum likelihood estimation (Koen et al., 2017) and the scores and standards are specifically determined for different populations and with careful using (of the tests and their scores) by teachers and students at classroom. We understand that a cut-off point could be estimated for a task or for a multiple battery of tasks and respecting the inter-individual characteristics of the examinees. However, the idea of cut score for a long battery might generate a bias that we can reduce by entering several variables such as age, gender and ethnicity (here considering the country of origin). These independent variables are useful during the analysis of moderation effect for any test such as the dichotic listening.

Concerning laterality, this is defined for one of two conditions, normally: right or left-handed. The laterality demonstrates the dominance of the hemisphere and that will have impact in the individual's performance (focusing the language competence). One of the tests or tasks well-known to measure the laterality is the dichotic listening test. Two stimuli are prompted in a evaluation setting: the participants should listening to words or texts simultaneously in both ears. The literature concludes the righ ear advantage considering the left hemisphere dominance (Hugdhal et al., 2008; Prete et al., 2018; Westerhausen et al., 2009).

Those variables are determinant to assign the relevance of items because each item and test will function differently according to the specificities of the individuals which will interfere in the testing results process. The specificity is related to the facets of testing and they are crucial to examine the minorities' context (Dornyei, 2014; Harsch & Rupp, 2011). Specifically, the laterality is focused in this short paper to understand the role of inter-hemispheric action (lateralization) in parallel with independent variables more commonly studied: chronological age, age of onset of acquisition (AOA), gender, nationality, socioeconomic status. The statistical analysis of significancy for the laterality will be ensured by the sensitivity-specificity values produced when examining the area under the curve (AUC) that the ROC analysis provide.

The ROC analysis allows a verification of prociency tests when they are administered to groups and subjects that may differ according to the variables mentioned above (Huibregtse et al., 2002). And the AUC concludes about the adequacy of tests for those groups and subjects. Regarding the administration of the tests we should outline: the training of raters/teachers to use correctly the tests and to calculate the scores based in the cut-off standard for specific populations (Harsch & Rupp, 2011; Hille & Cho, 2020; Wind & Peterson, 2018); the schedules of testing and the place where they are tested (at school, in their majority); and reduce the interference of the variation caused by the variables identified by the characteristics of the individuals (to apply statistics controlling for variables such as the laterality).

In the present study, auditory and visual tasks were administered but in this paper we are only addressing to the listening dichotic test in Portuguese as L2. Two different auditory stimuli were addressed by using computer and paper materials. The research question of this brief case report: can we consider the listening dichotic task as feasible measure of sensitivity and specificity to differentiate the school immigrants' performance?

2. Method

2.1. Participants

108 Portuguese immigrants identified as Second Language learners had a range of ages: between 7 and 17 years old and with origin in 20 countries. The participants, at the time of the study, lived in Portugal with their families and attended public schools. We had identified 21 mother tongues but the children spoke regularly above 30 languages at home, in total. Few are bilingual and also few students were correctly assessed concerning their L2 proficiency. This sample had, at the time of this report, the proficiency evaluated up to the B1 level (Common

European Framework of Reference for Languages, Council of Europe, 2001). All students attended different grades in school: first grade (1-4 years of schooling: 20.4%), second grade (4-6 years of schooling: 39.8%), third grade (7-9 years of schooling: 29.6%) and high school education (10-12th grades: 10.2%). They understood that the resolution of the tests was not part of the assessment weighing for regular school tests, being just used in a scientific study of the school immigrant population.

2.2. Instrument and Variables

The dependent variables of this study were: dichotic listening test with two inputs (4 words for each ear). The independent variables were gender, ethnicity, chronological age, age of onset of acquisition (age of acquisition of L2) and laterality.

The dichotic listening test is one of the battery of tests developed by the author and all them were computed to follow the recommended statistical procedure of ROC analysis. The individuals were asked to answer to a listening task (dichotic), by writing. That task was presented as a sound file with four words in each ear. After listening to the files simultaneously, the participants should report (by writing in a paper) as soon as possible how many words their heard. Among the eight words, four were non-words. Most of the non-words were registered by the examinees as Portuguese words (they retrieved words by visual similarity as a common cognitive strategy).

2.3. Procedures

The children's legal guardians filled out an informed consent prior to the start of the study. In the 2014-2017 period, students conducted the tests in a single session lasting approximately 60 minutes, in the same space that had been previously stipulated by the schools and on appropriately determined schedules. There were no absences among the 108 participants recruited for this sample. During the tests at each school, the room was prepared to avoid noise or distracting factors. In some cases, the researcher rescheduled, with the schools' authorization, a second session with the same students because some participants had not completed the task in the assigned session period considering that there were classes that only allowed, in specific school days, 45 minutes and not 60 minutes. The tests were done during the weekly school schedule. After participants entered the room that was established (by the school administration) for the research study, there was always a 10-minute preparatory moment when the researcher gave instructions on how to complete the tests and on the materials involved (paper and computer with headphones). The 1st Grade sessions took place mainly in the morning and the sessions of the remaining grades were completed in the afternoon.

Prior to administer the test in the schools, a pilot study was carried out with a sample of <35 subjects for linguistic correctness and to ensure that participants (considering the target group – immigrants in Portugal) would understand the questions, despite the proficiency limits. The order of tasks followed the type of linguistic levels underlying this diagnostic assessment: the lexical tasks were firstly grouped, followed by morphology and phonological awareness. Finally, there were tasks related to writing and listening skills. All data were first stored in the completed forms and in an electronic database - SPSS v. 26 - for statistical treatment and analysis, under complete data protection. It is important to outline that dichotic listening task was the main focus in this study and placed in the third group: phonological awareness. The data analysis followed the ROC curve analysis to examine the sensitivity and specificity scores. In the next section the ROC analysis is fully described.

3. Data Analysis, Results and Discussion

The ROC curve analysis was selected among another statistical measures attending to the inferential statistics previously used in robust literature addressing the cut scores for L2 and foreign language testing: determination of facets to identify the specificities of ethnic groups;

area under the curve (AUC) to identify what tests are adequate in a L2 testing and discrimination of learners (meaning: discrimination of proficiency levels where the participants should be placed).

Firstly, through the ROC analysis, we had proceeded to the baseline: the identification of the cut scores to find which participants are in different levels (satisfactory or non-satisfactory levels), then the specificity and sensitivity scores were calculated to determine the predictive value of the variables in this study.

The ROC analysis was administered in order to identify sensitivity and specificity scores and the cut-off point value. These scores will determine which participants are or are not at risk for Portuguese L2 learning. Considering the cut-off, children scoring higher than the cut-off point were placed in satisfactory proficiency level. This level did correspond to positive preparation for L2 learning, especially considering the lexical and semantic levels. The baseline for this study was to determine who are proficient and who are not.

Positive sensitivity values (differentiating those above and below the cut-off point) and also satisfactory specificity values (false positives) were calculated. It was found variation in scores according to variables such as laterality, gender and age. The laterality was our focus and the scores will be reported accordingly. With regard to laterality (right-handed/left-handed), all tests revealed a predictive value, however only for right-handed and not left-handed cases: *AUC: .60; cut-off point: 7.5 (8 words); sensitivity= .923; specificity= .842*. The values displayed are very high and the task presented a high level of difficulty considering the cut score approx. of 8 (the total number of words). The participants identified as right-handed has serious difficulties to attain the plenty level of 8 words decoded during the binaural test. To notice that the laterality is the independent variable related specifically with the dichotic listening task. The tests studied in this article did not differentiate, with significant values, left-handed immigrant students. So, we conclude that learners with more or less difficulty are right-handed. That differentiation is not applicable for the left-handed.

We effectively had identified satisfactory sensitivity and specificity indices which differentiated participants at the basic level of proficiency or without proficiency. Thus, the satisfactory cut-off values, through the ROC curve test, showed that listening dichotic test was suitable to integrate the diagnostic tool to be used with immigrant student population with age at 17 years old as maximum. If we observe the statistical data (cut-off of 7.5 and the AUC of .60) we can verify that this is a task of high difficulty for students' discrimination.

The identification of risk condition is useful to inform educational professionals about the level and difficulties that immigrant students may have through cut-off points determined and through variables such as laterality. In fact, educational professionals are not fully aware about the laterality/right-handed factor for the L2 learners' achievement (considering that millions of immigrants are immersed at classroom).

Being right-handed or left-handed will influence a different brain activation to process the same stimulus or the same test (Hiraoka et al., 2018; Krashen, 1973). However, the type of tasks in this study involves greater demand from the left areas of the brain, related to the understanding and coding of language in first or home language. The scenario appears to change in the case of which L2 is being used for the analysis. Here the L2 was the European Portuguese. Evidence such as semantics and morphology require the use of information stored in other brain areas than just the left. Older studies in the area of laterality and L2 have attested that the learners (mostly right-handed) coding in L1 activated the left areas more, whereas the subjects coding in L2 needed to activate right brain areas more. According to more recent studies, right-handed subjects, for example, in semantic decoding tasks experience greater difficulty precisely because they need greater use of non-dominant brain areas, needing to activate right areas for language (Bartha-Doering et al., 2018; Xue et al., 2004).

More studies should be developed for the dichotic listening task in other speakers and populations living as immigrants and mainly focusing the young school population. Additionally, is of upmost importance to replicate this study to understand the indices of sensitivity and specificity produced by the ROC analysis with a previous baseline established. Thus, the cut-off points for each task appear as definitive predictor to identify which tests we may use at schools and for specific groups of students (for example attending to laterality). Moreover, left-handed people should constitute in further analyses new samples to produce a comparative study with the right-handed. In fact, a limitation of the study was the population right-handed in high number by comparing to the left-handed subjects (in general studies). It would be important to advance a study with two groups more balanced concerning the right and left-handed. It is important to ultimate also that this piece of study did not observe influence of the controlled variables (gender, age) and the inter-individual variation is due to specific skills such as the listening discrimination. More studies should develop analysis with ROC curve and with binaural setting for assessment of L2 learners.

4. Acknowledgment

Acknowledgements: This work was funded by national funds through FCT - Fundação para a Ciência e a Tecnologia (Foundation for Science and Technology) - as part of the project CIP/UAL – Ref. UIDB/04345/2020; and the Psychology Research Centre (CIP) of Universidade Autónoma de Lisboa/Universidade do Algarve. Translation by Carolina Peralta.

References

Bartha-Doering, L., Kollndorfer, K., Kasprian, G., Novak, A., Schuler, A. L., Fischmeister, F. P. S., & Berl, M. M. (2018). Weaker semantic language lateralization associated with better semantic language performance in healthy right-handed children. *Brain and behavior, 8*(11), e01072. https://doi.org/10.1002/brb3.1072

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge University Press.

Dörnyei, Z. (2014). *The psychology of the language learner: Individual differences in second language acquisition*. Routledge. https://doi.org/10.4324/9781410613349

Fatahi, J., Tahaei, A. A., Ashayeri, H., Mohammadkhani, G., & Jalaie, S. (2018). Evaluation of Central Auditory Processing of Azeri-Persian Bilinguals Using Dichotic. I*ranian Journal of Child Neurology*, 13(1), 79-90.

Harsch, C., & Rupp, A. A. (2011). Designing and scaling level-specific writing tasks in alignment with the CEFR: A test-centered approach. *Language Assessment Quarterly, 8*(1), 1-33. https://doi.org/10.1080/15434303.2010.535575

Hille, K., & Cho, Y. (2020). Placement testing: One test, two tests, three tests? How many tests are sufficient? *Language Testing*. https://doi.org/10.1177/0265532220912412

Hiraoka, K., Igawa, K., Kashiwagi, M., Nakahara, C., Oshima, Y., & Takakura, Y. (2018). The laterality of stop and go processes of the motor response in left-handed and right-handed individuals. *Laterality: Asymmetries of Body, Brain and Cognition, 23*(1), 51-66. https://doi.org/10.1080/1357650X.2017.1311906

Hugdahl, K., Westerhausen, R., Alho, K., Medvedev, S., & Hämäläinen, H. (2008). The effect of stimulus intensity on the right ear advantage in dichotic listening. *Neuroscience letters, 431*(1), 90-94. https://doi.org/10.1016/j.neulet.2007.11.046

Koen, J. D., Barrett, F. S., Harlow, I. M., & Yonelinas, A. P. (2017). The ROC Toolbox: A toolbox for analyzing receiver-operating characteristics derived from confidence ratings. *Behavior research methods, 49*(4), 1399-1406. https://doi.org/10.3758/s13428-016-0796-z

Krashen, S. D. (1973). Lateralization, language learning, and the critical period: Some new evidence. *Language learning, 23*(1), 63-74. https://doi.org/10.1111/j.1467-1770.1973.tb00097.x

Kuiken, F., & Vedder, I. (2017). Functional adequacy in L2 writing: Towards a new rating scale. *Language Testing, 34*(3), 321-336. https://doi.org/10.1177/0265532216663991

Namaziandost, E., Abedi, P., & Nasri, M. (2019). The role of gender in the accuracy and fluency of Iranian upper-intermediate EFL learners' l2 oral productions. *Journal of Applied Linguistics and Language Research, 6*(3), 110-123.

Prete, G., D'Anselmo, A., Brancucci, A., & Tommasi, L. (2018). Evidence of a right ear advantage in the absence of auditory targets. *Scientific reports, 8*(1), 1-8. https://doi.org/10.1038/s41598-018-34086-3

Saxena, S., Ambler, G., Cole, T. J., & Majeed, A. (2004). Ethnic group differences in overweight and obese children and young people in England: cross sectional survey. *Archives of Disease in Childhood, 89*(1), 30-36.

Suzuki, S., & Kormos, J. (2019). Linguistic dimensions of comprehensibility and perceived fluency: An investigation of complexity, accuracy, and fluency in second language argumentative speech. *Studies in Second Language Acquisition*, 1-25. https://doi.org/10.1017/S0272263119000627

Ulbrich, C., & Wiese, R. (2018). Phonotactic principles and exposure in second language processing. *Empirical approaches to the phonological structure of words (Linguistische Arbeiten 567)*, 153-182. https://doi.org/10.1515/9783110542899-007

Wang, Y., & Wang, J. Q. (2002). A comparison of international references for the assessment of child and adolescent overweight and obesity in different populations. *European journal of clinical nutrition, 56*(10), 973. https://doi.org/10.1038/sj.ejcn.1601415

Wind, S. A., & Peterson, M. E. (2018). A systematic review of methods for evaluating rating quality in language assessment. *Language Testing, 35*(2), 161-192. https://doi.org/10.1177/0265532216686999

Westerhausen, R., Moosmann, M., Alho, K., Medvedev, S., Hämäläinen, H., & Hugdahl, K. (2009). Top-down and bottom-up interaction: manipulating the dichotic listening ear advantage. *Brain Research, 1250*, 183-189. https://doi.org/10.1016/j.brainres.2008.10.070

Xue, G., Dong, Q., Jin, Z., Zhang, L., & Wang, Y. (2004). An fMRI study with semantic access in low proficiency second language learners. *NeuroReport, 15*(5), 791-796. https://doi.org/10.1097/00001756-200404090-00010