ARTIFICIAL LANGUAGE TASKS AS A MEASURE TO ASSESS BILINGUALS' LEARNING SKILLS IN THIRD OR ADDITIONAL LANGUAGE ACQUISITION: THE CASE OF LLAMA_F

Francesca d'Angelo

HAL Id: hal-03216321

https://hal.science/hal-03216321

Submitted on 3 May 2021

ARTIFICIAL LANGUAGE TASKS AS A MEASURE TO ASSESS BILINGUALS' LEARNING SKILLS IN THIRD OR ADDITIONAL LANGUAGE ACQUISITION: THE CASE OF LLAMA_F

Francesca D'Angelo

University of Salerno (ITALY)
https://orcid.org/0000-0002-5428-8201

Abstract

The paper aims at discussing the validity and effectiveness of LLAMA_F, an artificial language test of grammatical inference, in research dealing with Third (or Additional) Language Acquisition (TLA). After describing the mechanism behind the test, the factors which make it particularly suitable to assess participants with a cognitive and linguistic background considerably different from second language learners are analysed. Indeed, to assess bilinguals' ability to learn additional languages, several problems arise, including level of competence reached in each language, age and order of acquisition, amount and type of exposure to the L3, learning context (implicit/explicit), typological proximity. Besides, the test is propounded as an effective and reliable tool allowing to assess Metalinguistic Awareness (MLA) and the transfer of learning strategies, developed by participants in previous languages. This is crucial to have an accurate portrait of fundamental factors proved to predict success in additional language learning, at any stage of educational contexts.

Keywords: Third Language Acquisition; Artificial Language Task; Bilingual Education; Metalinguistic Awareness; Cross-Linguistic Transfer.

1   INTRODUCTION

Third Language Acquisition (TLA) is a relatively new field of study, and area of research, developed considerably in the last few years. It refers to the study of a non-native language by learners who have previously acquired or are acquiring two languages. Specifically, it has been defined as 'the acquisition of a language that is different from the first and second and is acquired after them' (Cenoz 2013, 71). The study of TLA brings together two fields that have traditionally worked separately: i.e. second language acquisition (SLA) and the study of the effects of bilingualism on cognition and language learning.

Despite the similarities of SLA and TLA, there are several linguistic and cognitive factors to consider TLA as a distinct process and area of research from SLA. Second and third language learners come to the process of language acquisition with a linguistic and cognitive background that differs considerably, both quantitatively and qualitatively (Grosjean 1998). That is, a second language learner is a monolingual at the initial state of SLA, whereas a third language learner is already bilingual (potentially early/late, simultaneous/consecutive, etc.). This entails that having at least two languages in their linguistic repertoires allows third language learners to relate new structures, new vocabulary, or new ways of expressing communicative functions to the two languages they already know, not just one of them, as in the case of monolinguals. Moreover, third language learners show more refined skills and strategies for achieving the language-learning task (Garraffa et al. 2020).

Thus, if one takes into account the complexity of the field, due to the aforementioned quantitative and qualitative differences, the traditional assessing tools employed in SLA research may not be adequate. Indeed, not only does TLA differ considerably from SLA but the effect and interaction of the factors involved become difficult to observe and record experimentally. Specifically, to assess bilinguals' skills in TLA several problems arise: i.e. level of competence reached in each language, age and order of acquisition, exposure to the L3, learning context (implicit/explicit), typological proximity etc.

The present work aims at discussing the different advantages of using an Artificial Language Task, i.e. LLAMA_F test, in research dealing with multiple language acquisition. In particular, LLAMA_F is propounded as an effective technological tool to assess bilinguals' Metalinguistic Awareness (MLA), developed in previous languages, and to predict successful performance in TLA. It is a test of grammatical inference, which focuses on the participants' ability to learn additional languages in an informal setting. Loosely based on the Modern Language Aptitude Tests by Carrol & Sapon (1959), it

is part of the LLAMA battery developed by Meara (2005) as shorter, free, language-neutral tests. The first attempt at works in this area appeared as Meara, Milton, and Lorenzo-Dus (2002) and included a set of five tests assessing different aspects of language learning aptitude, i.e. vocabulary learning, grammatical inference, sound-symbol association, phonetic memory, and a test of memory for unusual sound sequences. The rise of interest generated in the research community, since the first publication, prompted the authors to adapt the tests for people with a different L1 from English. Moreover, another problem was that some of the material languages developed as part of the original test started to be familiar to potential participants. For instance, Polish and Turkish, although not widely recognised in the UK, are more familiar to test takers with Hungarian or Azeri as an L1. Hence, the need for a new set of tests, which was largely independent of the participant's L1, lead the designers to develop the actual version.

The main, practical reason making the test particularly suitable for research on bilingualism is that it is entirely based on picture stimuli and has eliminated the need for an L1 database. Second, not only does it allow to control for participants' L1, but it also allows to control for the amount of exposure and level of competence in L3 (i.e. the tested language). In other words, being an artificial language, test takers are not supposed to be familiar with the language of the task presented. Indeed, it would be difficult to find participants with the same level of proficiency in any natural language as an L3, same amount of exposure to the language, same method of acquisition and level of formal instruction received. Third, the task does not explicitly focus on grammar and MLA of the language as no instruction or request about the formal aspects of the grammar was included. Instead, as it will be better described in the following sections, the grammatical rules behind the artificial language are only implicitly inferred to match the sentence description with the picture. Additionally, the restricted time that participants have at their disposal to passively observe the unknown language and figure out the mechanism beyond it, as well as the fact that they are not allowed to take notes, make the task more similar to the implicit learning process of a language. Accordingly, the task does not directly facilitate those bilinguals with higher levels of MLA and language learning experience that are supposed to perform better in TLA.

2   METHODOLOGY

Participants are required to take the LLAMA_F, test of grammatical inference, as the last task of the session. The box marked   the grammar of the new of time available language, i.e. five minutes. The instructions provided are the following: "Your task is to use this time to learn as much as you can about a new language. You do this by clicking on the small buttons in the main panel. For each button you click, a picture and a sentence describing it will be displayed, as in the picture below" (see attachment 1).

Unak-ek eked-ilad is the sentence that describes the picture. The presentation phase of the program shows the test-taker a series of pictures depicting shapes and objects, and a short sentence in an artificial language that describes each image. Participants are expected to work out how the descriptions related to the pictures. From this, they should be able to intuit some of the grammatical and morphological features of the language such as word order, gender, singular, dual and plural numbers, conjugating prepositions etc.

During the second phase, they are presented with a new set of pictures that incorporate new elements as well as some taken from the training phase. Each picture is accompanied by two sentences that might describe it, and test-takers are required to select the correct description. They should be able to do this if they have internalised the grammatical rules evidenced in the presentation phase. Five points are awarded for a correct answer and five points deducted for an incorrect choice. One sentence is grammatically correct, while the other contains a major grammar error. The program gives them feedback in the form of a ding for a correct answer, and a bleep for an incorrect answer. To see the next test item, they need to click the   button. There the score, as they work through the test, and shows how many items are left to complete it. At the end of the test, the score, ranging between 0 and 100, is displayed on the bottom panel.

3   DISCUSSION

As already mentioned, LLAMA_F, test of grammatical inference, is entirely based on picture stimuli and has eliminated the need for an L1 database. This makes it an effective assessment tool for additional language learning as not only does it allow to control for participants' L1, but it also allows to control for other important variables, i.e. amount of exposure and level of competence in L3. Indeed, being an artificial language, test takers are not supposed to be familiar with the language of the task involved. In addition, the test resolves another important issue when recruiting participants in TLA studies. That is, the difficulty of finding a homogeneous sample of bilinguals with the same level of

proficiency in any natural language as an L3, the same amount of exposure to the language, the same method of acquisition and level of formal instruction received.

The second reason deals with a factor considered to be crucial in multilingual education, the relationship between Metalinguistic Awareness (MLA) developed in previous languages through formal instruction and level of attainment in TLA (D'Angelo 2020). Indeed, the test does not explicitly focus on grammar and MLA of the language as no instruction or request about the formal aspects of the grammar is included. In other words, the grammatical rules behind the unknown language are only implicitly inferred to match the sentence description with the picture. What is more, the restricted time that participants have at their disposal to passively observe the artificial language and figure out the mechanism behind it, together with the impossibility to take notes, make the task more similar to the implicit learning process of the language. Accordingly, it can be claimed that the task does not directly facilitate bilinguals with higher levels of MLA, supposed to perform better in TLA.

Moreover, the test allows assessing another crucial aspect of language learning, i.e. the ability to use and transfer the skills developed through bilinguals' previous experience as language learners, as it reproduces the initial stage of learning additional languages. Interestingly, the type of transfer does not apply to lexical items or grammatical structures from one language to the other(s), as suggested by the Typological Proximity Model (Rothman, 2011). Indeed, the performance in the additional language refers to an artificial grammar that is not typologically related to any of the languages already known by the test takers. Accordingly, the transfer refers to the practice of internalising grammatical patterns and exploiting those strategies when dealing with additional languages.

This is confirmed by the fact that, during the performance of the LLAMA_F, participants are required to make grammatical inferences on the basis of a short passive training where they are supposed to figure out how the language works. It is important to highlight that test-takers did not receive any explicit instructions about the grammatical rules and no linguistic terminology was employed. Therefore, what the test assesses in the second phase is precisely their ability to exploit their language learning strategies, grammatical inference, and their ability to generalise principles based on observation of the language by decoding and interpreting it. In other words, when learning another language, participants use their capacity to learn grammar through previous language learning rather than using their knowledge of any specific individual grammatical structures. Finally, as participants were not allowed to take notes during the training phase, the test requires additional effort involving the Working Memory as they had to remember the relevant patterns involved in the training to understand the mechanism of the language they were learning.

3.1 The validity of Artificial Language Tasks

Despite the acknowledged importance of ALL experiments as tools exploring the principles of language, as well as language learning ability, a persistent question is whether ALL studies can be considered as ecologically valid assessments of natural language ability. In particular, the aspect of artificial language tasks that have been questioned is that they cannot be compared to natural language learning. Nonetheless, several studies assessed the validity of artificial language tasks by comparing their performance with that recorded in natural language tasks, with a number of internal factors being controlled.

Ettlinger et al. (2015), for instance, bridged the gap between ALL and natural language learning research by comparing the performance of adult learners of Spanish as an L2 and ALL enhancement. The findings suggest that performance in ALL tasks correlates positively with indices of L2 learning even after controlling for IQ, general intelligence and the potential mediation of these internal factors. The authors also considered the effects of specific features of ALL tasks such as including or not a semantic aspect as well as presenting a complex or simple grammar. Interestingly enough, they inferred that ALL studies that incorporate a semantic component and involve more complicated grammatical systems may closely resemble the learning process of natural languages.

The specific validity of the LLAMA tests has been confirmed by a study by Rogers, Meara, and colleagues (2017) including 240 participants. In particular, the issues raised by the authors concern the following factors: whether the tests are language-neutral, the effects of bilingualism and age on LLAMA tests scores, and the amount of variance that background factors can account for in the LLAMA test results. Interestingly, the authors did not find any significant difference in terms of language background, suggesting that LLAMA tests are indeed language-neutral as there are no differences between groups once other factors (i.e. L2 instruction) are controlled for.

Concerning the effects of bilingualism, specifically, the difference referred to monolingualism, bilingualism, and instruction in L2, instructed L2 learners significantly outperformed the monolingual and bilingual group on two sub-components: LLAMA_B and LLAMA_F. The hypothesis of a training effect is also related to the results found in terms of the effects of age. Younger groups (10-11 years

old) were outperformed by older groups (20-21 Years old) suggesting that LLAMA tests are not suitable for children in that older learners may have developed more refined learning strategies over the years.

Finally, the effects of six background individual factors were also considered to see how much of the variance in the LLAMA test score could be accounted for by each of them: i.e. L1, L2 status, age, highest formal education, gender, and logic training (puzzles). The multiple regression results from 240 participants show that LLAMA tests can generally be used across different L1s, with male and female participants of different education levels and with different ages as these factors did not consistently affect the variance in scores. The only individual variable to predict most of the variance was prior instruction in a second language in LLAMA_B (6%) and LLAMA_F (2.6%). Therefore, from the analysis provided, it can be argued that LLAMA tests can be considered as robust and reliable measures of language learning ability.

4   CONCLUSIONS

The main focus of the work was to demonstrate the validity and effectiveness of LLAMA_F, test of grammatical inference, in multilingual research. Specifically, the test was used to overcome several issues due to the complexity and large amount of variation found in TLA research, being not subject to significant external factors or individual variables that could potentially affect the results. More specifically, the quantitative and qualitative complexity of the factors characterising studies on TLA, including number of languages, typological proximity, learning context, age and order of acquisition etc., requires different assessing instruments which allow to deal with the variability of the aforementioned factors as well as to take into account their interaction.

In terms of context of acquisition, it has been proved to be a valid tool to address the relationship between bilinguals' language learning experience, their explicit and implicit MLA, developed in other languages, and their ability to learn a third (or additional) language. Indeed, the ability to think about the language as an abstract system, made of different levels interacting among each other, and to focus on the grammatical form, to analyse and manipulate it, independently of the specific language involved, is hypothesised to be the most important predictor of success when learning additional languages. Bilinguals are also expected to enhance their explicit level of MLA in proportion to the amount of formal instruction they received. The number of languages mastered is another factor supposed to enhance MLA on the whole, which in turn assists the process of language learning.

As it has been highlighted, the input from the LLAMA_F is unrelated to the linguistic structure of any other potential language known by the test takers. It allows inferring that it is not the specific grammatical knowledge of a particular language that boosts the process of TLA. Instead, it is the abstract knowledge of the language as a system that assists participants to think about the language critically, as an abstract object, and to resort to this awareness and ability when dealing with additional language learning. In other words, when they come to the process of third or additional language acquisition, they do not need to relearn the fundamental principles of language structures. They make use of these explicit and abstract fundamental rules to figure out how the new language works by applying those principles to additional languages.

Thus, the implication of this condition of MLA enhancing the process of additional language learning in bilinguals is that the content of metalinguistic knowledge must be broader than any that applies to the knowledge of a specific linguistic structure. In this perspective, MLA refers to the 'explicit representation of abstract aspects of linguistic structure that become accessible through the knowledge of a particular language' (Bialystok 2001, 124). Specifically, in TLA research, the aforementioned knowledge is reinforced through the knowledge of at least two other languages as it deals with bilinguals' additional language acquisition and has proved to be the most important predictor of language attainment.

Accordingly, the notion of transfer of the language skills from the first to the second language if there is sufficient exposure to the L2 and motivation to learn the language can be applied by extension to TLA. That is to say, the language skills developed in an L2, in a broader and more abstract knowledge, will be transferred to the L3. In line with this hypothesis, the test confirms that the level of attainment reached in the third language is affected by participants' ability to manipulate, analyse, and think about the L2 as an abstract object.

The empirical evidence that the more languages bilinguals have gained literacy and study experience in, the better they are at learning additional languages with an implicit focus on grammatical form confirms the claim supported by many scholars in the field. As Cenoz (2013) points out, MLA is one of the key factors associated with bilinguals' better performance in TLA together with learning strategies and a broader linguistic repertoire. These three elements affect each other and are closely related to the number of languages known by bilingual learners. First, the higher level of MLA can be considered

as both a cause and effect of bilinguals' success in additional language learning. On one hand, based on their previous experience of the task of learning a language and their knowledge of two linguistic systems they enhance their level of MLA. On the other, it is precisely the more developed MLA skills and the idea that they manage to think about the language in an abstract way and regard it as an object to assist them in the process of learning additional languages.

Indeed, the claimed bilingual advantage in TLA is related to bilinguals' experience as language learners itself as they develop a wider range of learning strategies. Specifically, it has been argued that they look for more sources of input, make an early effort to use the new language, and show self-direction and a positive attitude towards the task (Bowden et al. 2005). This argument, mostly developed on the basis of research using artificial language tasks, has also been confirmed by other studies dealing with natural languages.

Also, the broader linguistic repertoire that bilinguals have at their disposal has been associated with a better performance in TLA. The majority of the studies have explained it in terms of language distance, that is, closely related languages would be more useful for bilinguals learning a third (e.g. De Angelis 2007, Ringbom 2007, Rothman, 2011). However, as already mentioned, this is not the case for LLAMA_F, as the task has been proved to be language-neutral and to have no typological relationship with the other languages known by the participants. This suggests that the broader linguistic repertoire on behalf of bilinguals supports them in TLA independently of nature and specific linguistic features of the languages involved.

To sum up, for all the reasons discussed in the present paper, it can be claimed that LLAMA_F represents a valid and reliable technological instrument allowing to overcome several specific issues when dealing with TLA research. In particular, what the test assesses is precisely the participants' ability to exploit their language learning strategies, grammatical inference, and their ability to generalise principles based on the observation of the language by decoding and interpreting it. In other words, when learning another language, participants use their capacity to learn grammar through previous language learning rather than using their knowledge of individual grammatical structures.

What is more, in the recruitment phase, it eliminates the need for an L1 database, being entirely based on picture stimuli. Also, it permits to control for the amount and type of exposure to the L3, since test takers are not supposed to be familiar with the artificial language presented. For the same reason, it does not affect the typological relation with any previous language known by participants since, as discussed, the test has been proved to be "language neutral". Finally, the test reproduces the implicit language learning process and does not facilitate those participants with a higher level of explicit knowledge of the language, allowing to test the effects of MLA, i.e. a crucial variable affecting the process and outcome of additional language learning.

REFERENCES

Bialystok, E. (2001) *Bilingualism in Development: Language, Literacy and Cognition.* Cambridge, Cambridge University Press.

Bowden, H., Sanz, C., Stafford, A. (2005) 'Individual Differences: Age, Sex, Working Memory, and Prior Knowledge' in C. Sanz (Ed.), *Mind and Context in Adult Second Language Acquisition: Methods, Theory and Practice.* Washington DC, Georgetown University Press: 105-140.

Carroll, J. B., Sapon, S. M. (1959) *Modern language aptitude test.* San Antonio, TX, US, Psychological Corporation.

Cenoz J. (2013) 'Defining Multilingualism', *Annual Review of Applied Linguistics*, 33, pp. 3-18.

D'Angelo, F. (2020) 'The Additive Effect of Bilinguals' Metalinguistic Awareness in Additional Language Acquisition', In: Patrick-André Mather, (eds.), *Technology-enhanced Learning and Linguistic Diversity: Strategies and Approaches to Teaching Students in a 2nd or 3rd Language* (Emerald Studies in Higher Education, Innovation and Technology), 13-24. Bristol: Emerald Publishing.

De Angelis J. (2007) *Third or Additional Language Acquisition.* Clevedon, Multilingual Matters.

Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., Wong, P. C. (2015) 'The Relationship Between Artificial and Second Language Learning', *Cognitive science*, 40(4), pp. 822-47.

Garraffa, M., Sorace, A., Vender, M. (2020) *Il cervello bilingue.* Roma: Carocci editore.

Grosjean F. (1998) 'The on-line processing of speech: Lexical access in bilinguals', in P. Bhatt, R. Davis, (Eds.), *The Linguistic Brain.* Toronto: Canadian Scholars' Press.

Meara, P. (2005) *LLAMA Language Aptitude Tests: The Manual* (Tech. Rep.). Swansea: Lognostics.

Meara, P., Milton, J., Lorenzo-Dus, N. (2002) *Swansea language aptitude tests* (LAT) v2.0. Swansea: Lognostics.

Ringbom H. (2007), *Cross-linguistic Similarity in Foreign Language Learning*. Bristol: Multilingual Matters.

Rogers, V., Meara, P., Barnett-Legh, T., Curry, C., Davie, E. (2017) 'Examining the LLAMA aptitude tests', *Journal of the European Second Language Association*, 1(1), pp. 49-60.

Rothman J. (2011) 'L3 syntactic transfer selectivity and typological determinacy: The typological Primacy Model', *Second Language Research*, 27(1), pp. 107-127.

Attachment 1.

LLAMA_F: an example of picture stimuli used in the presentation phase.