Learning and Individual Differences 50 (2016) 42–48

Contents lists available at ScienceDirect

Learning and Individual Differences

journal homepage: www.elsevier.com/locate/lindif

The LLAMA tests and the underlying structure of language aptitude at two levels of foreign language proficiency

Gemma Artieda *, Carmen Muñoz

University of Barcelona, Spain

ARTICLE INFO

Article history:
Received 6 September 2015
Received in revised form 23 April 2016
Accepted 24 June 2016
Available online xxxx

Keywords:
Language aptitude
Individual differences
LLAMA tests
Foreign language acquisition
Implicit learning

ABSTRACT

Language aptitude is one of the major individual differences under scrutiny in a study investigating the differential role of individual differences in foreign language learning at two levels of proficiency. Participants are adult Catalan-Spanish bilingual students of English. They were divided into a group of beginners ($N = 52$) and a group of intermediate learners ($N = 88$). Findings on language aptitude as measured by the LLAMA tests and expressed as separate aptitude components provide support for the differential impact of individual differences at two levels of proficiency. The results indicate that on a componential level phonemic coding ability impacts learners at the early stages of second language development, while language analytic abilities have an impact on learners at all levels. Finally, findings suggest that implicit learning processes seem to play a role in the early stages of adult foreign language acquisition.

1. Introduction

1.1. Language aptitude

In individual differences research, language aptitude and motivation are the two individual differences (IDs) which predict second language acquisition success more consistently, with correlation levels only equalled by age of onset (Dörnyei & Skehan, 2003). Correlations between language aptitude or motivation and second language (L2) proficiency are usually in the range of 0.20 to 0.60 (Dörnyei & Skehan, 2003), making language aptitude the most important cognitive individual difference affecting L2 acquisition.

Language aptitude has been defined as 'a range of different cognitive factors making up a composite measure that can, in turn, be referred to as the learner's overall capacity to master a foreign language' (Dörney, 2005:249). Recently, this definition has been supplemented by Robinson (2013), who highlights the pervasiveness of language aptitude effects across learning contexts:

> 'higher aptitude […] predicts more successful adaptation to instructive or naturalistic exposure to the L2, as measurable by demonstrable faster progress in learning, and in higher levels of ultimate attainment in proficiency at the end of a course of instruction, or following a period of naturalistic exposure to the L2' (Robinson, 2013:1).

The study of foreign language (FL) aptitude has been closely linked to the development of FL tests since its inception. In the early times, when prediction went ahead of theory building, Carroll (1981) proposed his four-factor structure of language aptitude based on the results of factor analyses of a large number of individual learner characteristics expected to contribute to FL learning. Empirically derived, Carroll's four factors were: phonemic coding ability, the ability to discriminate and code unfamiliar sounds in such a way that they can be recalled later; associative memory, the ability to make connections between native language words and their FL equivalents; grammatical sensitivity, the ability to identify the functions of words in sentences; and, finally, inductive language analytic ability, the successful identification and extrapolation of patterns between form and meaning.

Skehan (1998) set out to update the construct of language aptitude and proposed a three-component structure. The first component was phonemic coding ability, the same factor that Carroll had proposed. The second component was language analytic ability, which subsumed Carroll's grammatical sensitivity and inductive language learning analytic ability, and Skehan (1998:203) defined it as the ability to recognise patterns in language and to extrapolate and produce new chunks of language from the newly internalised rules. The third factor was memory, which Carroll had explained as associative memory in line with memory research in his times, but that for Skehan needed the ability to 'retrieve it efficiently in real time to handle natural conversational

* Corresponding author.
E-mail address: gartieda@ub.edu (G. Artieda).

http://dx.doi.org/10.1016/j.lindif.2016.06.023
1041-6080/

demands' (p. 204), a memory concept which is closer to the present understanding of working memory.

Drawing from the evidence gathered from successful and unsuccessful learners, Skehan (1998) proposed the following relationships between language aptitude components and L2 proficiency: phonemic coding ability is of greatest importance at the early stages of L2 proficiency, and it plateaus after that; after a threshold has been reached, the contribution of phonemic coding ability to L2 proficiency decreases substantially; language analytic ability has a linear relationship with success at all stages of the proficiency ladder, and it is therefore equally important at all levels; finally, memory is a fundamental component which is equally important at all levels of L2 development until an advanced level of proficiency is reached. At that point its importance increases and it becomes the determining factor for learners to achieve native-like command of the language.

In a similar line, Robinson (2005) proposed clusters of abilities which may be important at beginning, intermediate and advanced levels of L2 development. Ten basic cognitive abilities would contribute to input processability in early stages of L2 learning: processing speed, pattern recognition, phonological working memory capacity, phonological working memory speed, semantic priming, lexical inferencing, text working memory capacity, text working memory speed, grammatical sensitivity, and rote memory. Other abilities and traits would be necessary in advanced levels of L2 development, such as interactional intelligence, openness to experience, and pragmatic ability, all of these traits contributing to information processes and mediated by the demands of the tasks. Robinson (2005, 2013) has argued that to date no aptitude test takes a developmental approach to language aptitude by aiming at tapping at the different aptitude components which may play a role at different stages of L2 development.

Recently, Doughty et al. (2010) have been investigating the factor structure of a new aptitude battery, high-LAB, a new measure to identify individuals who have the aptitude to reach high levels of L2 proficiency, and with the potential to identify other aptitude components that may be key in later stages of L2 proficiency. In a more recent study coming from the same research project, Linck et al. (2013) have reported that high-level attainment is significantly related to working memory, explicit associative learning, and implicit learning.

A number of studies have aimed at testing Skehan's contention that language aptitude is multi componential. Sparks, Patton, Ganschow, and Humbach (2011) carried out a factor analysis of a test battery which included measures of language skills in the student's first language (L1) and in the L2, in a sample consisting of 54 high-school students. Two of the four factors identified in the factor analysis incorporated similar measures of L1 and L2 skills: language analysis, including vocabulary, language comprehension, grammar, and inductive language learning; and phonology/orthography, which used measures of phonetic coding and phonological processing (word decoding, spelling, and sound-symbol correspondence). This study provided evidence for the componential nature of language aptitude across languages (Skehan, 1989; Robinson, 2005, 2013), as well as emphasizing the long-term relationships between L1 and L2 learning. Results also supported the linguistic coding differences hypothesis' tenet that a phonemic coding deficit or poor language analysis in the L1 will be reflected on a student's L2 learning skill (Sparks & Ganschow, 1991, 1993, 1995; Sparks, Javorksky, Patton, & Ganschow, 1998).

A study by Harley and Hart (1997) differed from those above in that it had the specific aim of finding out whether early immersion in an L2 would enhance language aptitude. The comparison of a group of early immersion students and a group of late immersion students (a total of 65 11th grade students in a Canadian French immersion setting) did not support their hypothesis, but it showed that memory skills were more strongly associated with outcomes in the early immersion group, whereas analytical language ability was a better predictor of outcomes in the late immersion group. Thus, the study seemed to indicate that different components of aptitude may play a different role at different ages when interacting with different age-related instruction conditions.

1.2. Research on the componential nature of language aptitude using the LLAMA tests

Recently, the claim that language aptitude is a multifaceted construct resulting in L2 aptitude profiles (Skehan, 1998, 2002, 2012; Dörnyei, 2005) or L2 aptitude complexes (Robinson, 2005) has merited researchers' attention. This is the case of Granena (2011, 2012), who investigated whether the LLAMA subtests measured a unitary trait, labelled as language aptitude, or multiple aptitude subcomponents.

The LLAMA battery of tests which is used nowadays is an updated version of the initial exploratory suite (Meara, Milton, & Lorenzo-Dus, 2001), though Meara (2005) cautions researchers that it should not be used in high-stakes situations because it has not been standardized or validated. Recently, Granena (2013) conducted an exploratory validation study of LLAMA using a 186 participant sample from three different language backgrounds (English, Spanish, and Chinese). Results yielded acceptable levels of reliability, approaching an internal consistency coefficient of 0.80, as well as showing stability on a test-retest reliability procedure.

The LLAMA suite consists of the following tests:

LLAMA B: A vocabulary learning task aimed at measuring the ability to learn large amounts of vocabulary in a short space of time. Similar to Carroll and Sapon's vocabulary learning task, this version is language-independent as it uses visual stimuli rather than text.

LLAMA D: A task to measure how effectively the participant can recognise short segments of oral language to which they have been exposed previously. Unlike LLAMA B, LLAMA D is not based on Carroll and Sapon's work. It is inspired by the research by Service (1992) and Speciale (Speciale, Ellis, & Bywater, 2004), who claim that the ability to recognise patterns in oral language is a key skill for language learning. The sound sequences are computer generated and based on a dialect of an Indian language spoken in British Columbia (Canada). Besides, the spoken language has been synthesised using the AT&T Natural Voices for French, to make the sounds even more difficult to recognise by test takers.

LLAMA E: This sound-symbol correspondence task consists of a set of 22 recorded syllables which the participant needs to match to a transliteration of the syllable sounds in an unfamiliar language.

LLAMA F: A grammatical inferencing task that uses visual stimuli, thus making the test language-independent and usable by participants with any L1. Due to the limitations of using visual stimuli, LLAMA F relies more on agreement features than on word order.

In her literature review of 11 studies that used the LLAMA tests to measure language aptitude (Abrahamsson & Hyltenstam, 2008; Forsberg & Sandgreen, 2013; Granena, 2012; Granena & Long, 2013; Bylund, Abrahamsson, & Hyltenstam, 2010; Cherciov, 2011; Serrano & Llanes, 2012); Yalcin, 2012; Yilmaz, 2013; Smeds, 2012), Granena concluded that LLAMA results 'correlated with L2 measures that call for the use of analytic, metalinguistic abilities and with L2 learning under explicit instructional treatments or feedback conditions' (2013:112). Granena highlighted that only the LLAMA D sound-recognition subtest showed a significant relationship with L2 measures that required automatic use of the L2 (e.g. spontaneous oral production tasks).

Granena (2013) performed two exploratory Principal Components Analyses (PCAs): a first PCA resulted in a two-factor solution: LLAMA B, E, and F subtests (vocabulary learning, sound-symbol association, and grammatical inferencing, respectively) loaded on one factor, while LLAMA D (sound recognition) loaded on a second factor. A second PCA was then performed on a battery of six cognitive tests, including the