Language Awareness

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rmla20

Are individual differences in cognitive abilities and stylistic preferences related to multilingual adults' performance in explicit learning conditions?

Karen Roehr-Brackin, Gabriela Adela Gánem-Gutiérrez, Lexa Olivera-Smith & María Teresa Torres-Marín

To cite this article: Karen Roehr-Brackin, Gabriela Adela Gánem-Gutiérrez, Lexa Olivera-Smith & María Teresa Torres-Marín (2021): Are individual differences in cognitive abilities and stylistic preferences related to multilingual adults' performance in explicit learning conditions?, Language Awareness, DOI: 10.1080/09658416.2021.1969404

To link to this article: https://doi.org/10.1080/09658416.2021.1969404

Published online: 09 Sep 2021.

LANGUAGE AWARENESS
https://doi.org/10.1080/09658416.2021.1969404

Routledge
Taylor & Francis Group

Are individual differences in cognitive abilities and stylistic preferences related to multilingual adults' performance in explicit learning conditions?

Karen Roehr-Brackin, Gabriela Adela Gánem-Gutiérrez, Lexa Olivera-Smith and María Teresa Torres-Marín

Department of Language and Linguistics, University of Essex, Colchester, UK

ABSTRACT

Research suggests that individual differences in additional language learning may play a more important role in taxing situations when learners are confronted with unfamiliar or difficult tasks. However, studies to date have mostly focused on second language (L2) learners/bilinguals, while individual differences within multilingual populations remain under-researched. Working with university-level multilingual adults, we compared the effectiveness of traditional instruction (familiar) and concept-based instruction (unfamiliar) to teach the past tense/aspect distinction in Spanish. Learners were pre- and post-tested on their knowledge of the target structure and assessed on language learning aptitude, working memory capacity, verbal-imagery cognitive style and attitudes. While both treatment groups demonstrated significantly improved metalinguistic knowledge, we found no statistical differences between the two groups in terms of knowledge gained or attitudes, and individual differences in cognitive ability were not associated with observed gains. This set of results indicates that the cognitive individual differences measured seemingly no longer played a significant role in these multilinguals' performance in the instructional conditions examined. In addition, neither language learning experience nor typological closeness between known lan guages had any significant impact. We propose that extensive expe rience with explicit language instruction may have led to a levelling effect, as previously observed in L2 learners.

ARTICLE HISTORY
Received 27 September 2019
Accepted 9 August 2021

KEYWORDS
Multilingualism; second language (L2) learning; explicit instruction; concept-based instruction; Spanish past tense/aspect; individual learner differences

Introduction and background

Multilingual language learning

The notion of multilingualism is notoriously hard to define. In the present paper, we follow a recent call to conceptualise multilingualism as well as the related notions of bilingualism and of language itself as natural categories situated in specific cultural contexts (Berthele, 2021). Natural categories are radial, graded and have fuzzy boundaries (Rosch, 1978; Rosch & Mervis, 1975), so they cannot be defined with reference to necessary and sufficient features

CONTACT Karen Roehr-Brackin kroehr@essex.ac.uk

in the Aristotelian tradition (Roehr, 2008). Natural categories have prototypical members at their centre and more peripheral members closer to the category boundaries. According to this point of view, a language is a set of linguistic features used in speech and writing that is perceived to be different from neighbouring sets of features and for which a commonly agreed label is used. Language may thus be characterised as a natural category that varies along the dimensions of (a) recognition by the speech community and (b) perceived linguistic distance to neighbouring varieties (Berthele, 2021). To exemplify, in such a conceptualisation, Basque would be considered a prototypical language because it is both maximally distant from its neighbours and fully recognised by speech communities as a separate language. Examples of regular languages would be Javanese or Dutch. Javanese has a large distance to its neighbours, but is less recognised as distinct by (some) speech communities. Conversely, Dutch is fully recognised as distinct, but it has a smaller perceived distance to neighbouring languages such as German, for instance. Ndau and Luxemburgish can be considered examples of limiting cases that are close to the category boundaries, the former due to its low recognition as a separate language, the latter due to the minimal distance to its neighbours. Michigan English is cited as an example of a construct outside the category boundary, since it is neither sufficiently distant from other varieties of English nor recognised by any speech community as a separate language (Berthele, 2021, p. 102).

If we likewise conceptualise bi- and multilingualism as natural categories, a prototypical bilingual will be a person who regularly uses two prototypical languages at high levels of proficiency. If a distinction between bi- and multilingualism is considered useful, then a prototypical multilingual's repertoire includes more than two such languages. A regular multilingual may have one dominant and two non-dominant languages, or slightly skewed dominance in two languages and one dialect. An example of a limiting case of multilingualism would be a speaker with excellent command of various styles and registers in just a single language (Berthele, 2021, p. 105). In our subsequent discussion of existing research, we refer to category prototypes when we mention languages, bilingualism and multilingualism.

Recent work in additional language learning has increasingly emphasised not only quantitative, but also qualitative differences between multilinguals and bilinguals, or between second language (L2) learners and learners of a third (L3) or further (Lx) languages (Cenoz & Jessner, 2009; Jessner, 1999, 2006). This qualitative difference is conceptualised in terms of the so-called M-factor (Cenoz & Jessner, 2009; Jessner, 2008, 2014), with 'M' standing for multilingualism.

It has been suggested that both bilinguals and multilinguals have enhanced metalinguistic abilities, which draw on cross-linguistic awareness or "the learner's tacit and explicit awareness of the links between their language systems" (Cenoz & Jessner, 2009, p. 127). Metalinguistic abilities comprise the ability to separate language form and function, or to look through language meaning at language form. They also include the ability to treat language as an object of inspection, reflection and analysis (Baker, 2006; Bialystok, 2001; Gombert, 1992). Importantly, similarities between bi- and multilinguals are complemented by a number of proposed differences that are attributable to different prior language learning experiences. Specifically, it has been argued that bilingualism or L2 learning relates to a monolingual system as the norm, whereas multilingualism or L3/Lx learning relates to a bilingual system as the norm (Jessner, 1999, 2006, 2008, 2014). In an instructed context, L3 learning builds on L2 learning and thus shares certain characteristics with L2 learning.

However, L3 learning is also influenced by the degree of bilingualism the learner has already achieved.

It has been argued that multilingual learning is by definition more complex than bilingual learning (Cenoz & Jessner, 2009; Hufeisen, 2018). In the case of bilingualism, the two languages involved may either be learned simultaneously or consecutively, but in the case of multilingualism, possibilities multiply. Simultaneous trilingualism, consecutive learning of the first language (L1), L2 and L3, simultaneous learning of L2 and L3 after L1, and simultaneous acquisition of L1 and L2 prior to L3 are all possible scenarios. The learning process of a specific language may also be temporarily interrupted when another language is added (Cenoz & Jessner, 2009). Thus, the M-factor comprises everything that distinguishes a multilingual from a monolingual system. It refers to qualities that develop in multilinguals due to contact and experience with different languages, such as formal language learning experience, language management and language maintenance skills, and, last but not least, metalinguistic awareness and abilities. These can have a catalytic or accelerating effect in L3/Lx learning that is not necessarily observable in L2 learning (Jessner, 2008, 2014).

In keeping with this point of view, several studies have compared multilinguals not just with monolinguals, but also with bilinguals on variables such as metalinguistic awareness, language-analytic ability or working memory (e.g. Biedroń & Szczepaniak, 2012; Poarch & van Hell, 2012). By contrast, comparisons within multilingual populations are rarely made (Ortega et al., 2016). Yet there is no reason to assume that individual learner differences play a lesser role in multilinguals than in bilinguals or monolinguals. Indeed, it could be hypothesised that individual differences are potentially *more* important in such populations, given that multilingual language learning and use is characterised by greater complexity than bilingual language learning and use.

Individual learner differences

Individual differences refer to attributes which all learners possess, but on which individual learners differ (Dörnyei, 2005). These attributes include affective, conative, personality and cognitive variables. Focusing on the latter two categories, language learning aptitude, working memory (WM) capacity and cognitive or learning style are the most prominent factors. (Dörnyei, 2005).

It is worth noting that investigations into the role of cognitive style in L2 learning are comparatively rare (but see Littlemore, 2001; Ziętek & Roehr, 2011), even though theoretical arguments for the potential impact of learners' stylistic preferences have been put forward (Roehr, 2008; Sternberg & Grigorenko, 1997). Conversely, research to date has yielded strong evidence for the role of language learning aptitude in additional language learning (Wen et al., 2017). More than 60 empirical studies investigating the role of aptitude in L2 learning have been conducted over the past decades, and two recent meta-analyses of this research (Li, 2015, 2016) provide a concise and informative overview of the cumulative results.

Li (2015) was concerned with the relationship between aptitude and morphosyntactic attainment. The reported meta-analysis yielded an aggregated effect size of .31 with narrow confidence intervals, thus suggesting a robust effect. In other words, aptitude explained about 10% of the variance in morphosyntactic learning, which is quite considerable if it is borne in mind that the L2 learning process is influenced by a range of learner-internal and learner-external factors. Interestingly, it was also found that aptitude affected high school

students' learning more than university students' learning. As high school students are likely to be less advanced L2 learners on average, it appears to be the case that aptitude plays a more important role at lower proficiency levels. In addition, university students are a select group who may be more homogeneous and advanced, leading to weaker or no correlations (Li, 2015).

Li (2016) examined the relationship between aptitude and general L2 proficiency attainment and uncovered a correlation coefficient of .49. This indicates that aptitude was a strong predictor, with about 25% of the variance in general proficiency accounted for. Again, the mean effect size was larger for high-school than for university-level learners, once more suggesting that aptitude may be a better predictor at lower levels of proficiency, if one assumes that high school learners have had less L2 experience than university learners. The previously mentioned alternative explanation of more homogeneous scores among the self-selected population of university-level learners attenuating any effects still applies as well. Taken together, the findings from these two recent meta-analyses offer powerful evidence for a significant role of aptitude in L2 learning, with aptitude scores significantly predicting attainment in terms of both general proficiency and knowledge of morphosyntax.

Research to date has likewise uncovered evidence for the role of WM capacity in L2 learning (e.g., Juffs & Harrington, 2011; Linck et al., 2014; Wen, 2012), although results are overall less consistent and more complex than in the case of aptitude. This is at least in part due to the fact that there are many more measures of WM than measures of aptitude available, so the methods used vary across studies. The general hypothesis is that greater WM capacity will lead to more successful L2 learning, but results seemingly depend on whether researchers use measures of complex WM tapping both storage and processing, or measures of phonological short-term memory only (Juffs & Harrington, 2011).

In addition to the measures of WM that are used, the criterion measures also have a role to play. In other words, effects of individual differences in WM may be task-dependent. Furthermore, learners' level of L2 proficiency is likely to mediate the predictive power of WM. Thus, research in the area of vocabulary learning suggests that the role of phonological short-term memory is greater at lower levels of proficiency. With regard to general L2 proficiency, cumulative results point towards a weak contribution of phonological short-term memory in instructed L2 learning and a somewhat greater role for complex WM in the context of tasks requiring focused attention, noticing and explicit knowledge (Juffs & Harrington, 2011).

Instructional approach, proficiency level and learning difficulty

It goes without saying that L2 learning does not happen in a vacuum, so a single predictor can never explain all of the observed variance in outcomes. In multilingual learning in particular, it is evident that a multitude of variables interacts, so the role of individual differences in cognitive ability can be expected to interact with other learner-internal and learner-external factors, including a learner's level of proficiency and language learning experience (Rodríguez Silva & Roehr-Brackin, 2016; Serafini & Sanz, 2016), the type of instruction they are exposed to (Goo et al., 2015; Norris & Ortega, 2001), and the learning difficulty of the linguistic target they are attempting to master (Tomak & Roehr-Brackin, 2017).

The general question of which type of instruction might work best for adolescent and adult learners has been investigated in a considerable number of studies over the past decades. Both classroom-based and laboratory-based research has compared the relative effectiveness of various explicit and implicit teaching and learning conditions. The cumulative findings from (quasi-)experimental studies conducted over a period of more than 30 years have been reported in two meta-analyses (Goo et al., 2015; Norris & Ortega, 2001), which draw together the findings of 72 primary studies conducted between 1980 and 2011. An instructional approach can be defined as explicit "if rule explanation comprised any part of the instruction (…) or if learners were directly asked to attend to particular forms and to try to arrive at metalinguistic generalizations on their own", whereas an approach is considered implicit "when neither rule explanation nor directions to attend to particular forms were part of the treatment" (Norris & Ortega, 2001, p. 167). Based on these definitions, both meta-analyses confirm that explicit, form-focused instruction can effectively enhance L2 learning (Goo et al., 2015; Norris & Ortega, 2001).

Needless to say, these meta-analytic results are based on group means and are intended to reveal general cumulative effects; they do not distinguish between learners with different individual profiles or between linguistic targets of differing learning difficulty. Several empirical studies have provided supporting evidence for the interaction between individual learner differences on the one hand and learning difficulty, proficiency level and instructional approach on the other hand, though no study to date has included all these factors in a single research design.

In an early study investigating the association between language learning aptitude and instruction, Wesche (1981) worked with Canadian public service workers who were learning either English or French in a large-scale government training programme. Following the creation of detailed individual profiles based on data from an aptitude test, a learning style questionnaire, and an interview, learners were matched with one of three instructional approaches intended to cater for their respective strengths. Most learners had experienced the default audio-visual approach, but following learner complaints about that method, some were allocated to newly formed deductive-analytical and functional-situational streams, in accordance with their individual profiles.

Overall, the researcher found that learners who had been allocated to the deductive-analytical stream had more positive attitudes now that the teaching method matched their profile. Moreover, learners experiencing that approach achieved higher scores on three of four achievement measures that were used to assess listening comprehension and oral expression. For the functional-situational approach, only anecdotal evidence was available, but that evidence pointed in the same direction. In other words, offering learners a teaching approach that appeals to their specific cognitive strengths seems to result in improved outcomes and greater satisfaction (Wesche, 1981).

Later, Erlam (2005) found that language-analytic ability as a component of aptitude was particularly beneficial in an inductive instructional condition, and both language-analytic ability and WM capacity played a role in an instructional condition relying on structured input. Conversely, individual differences in aptitude and WM did not have any impact in a deductive instructional condition, which was closest to the type of instruction the participating English-speaking teenage learners normally experienced. The linguistic target in the study was French direct object pronouns.

Tomak and Roehr-Brackin (2017) reported that complex WM was a significant predictor of teenage learners' performance on oral measures of L2 achievement in an implicit,

meaning-focused instructional condition that encouraged incidental learning, but not in two explicit instructional conditions that encouraged learners to focus on form. In the explicit condition that made use of deductive instruction, aptitude predicted learners' performance on the oral measures, but not on the written measures used. Nine different uses of English articles constituted the linguistic target, which was considered a difficult structure for the participating Russian learners whose L1 does not instantiate articles.

Rodríguez Silva and Roehr-Brackin (2016) targeted 13 different linguistic structures of English with varying levels of difficulty. Individual differences in complex WM and, marginally, in aptitude predicted Spanish-speaking adult learners' performance on oral measures at the lowest (intermediate) proficiency level. In the two higher (upper intermediate) levels, or indeed on written measures at any level, differences in cognitive abilities no longer played a significant role.

Serafini and Sanz' (2016), study provides further evidence that WM is more important at lower proficiency levels. Working with adult English-speaking learners of Spanish, the researchers found that complex WM and phonological short-term memory significantly facilitated L2 development in learners who had had less exposure to Spanish, while this effect was not observed in the more advanced learners. The researchers suggest that exposure to explicit instruction over time may have levelled the playing field, making individual differences in cognitive capacity less relevant at more advanced levels of proficiency.

Research issues

In summary, research to date suggests that multilinguals differ from monolinguals primarily due to the so-called M-factor, which emphasises the enhanced metalinguistic abilities of L3/Lx learners. Nevertheless, individual learner differences can be expected to play a role, i.e. multilinguals are a no more homogeneous group of people than bilinguals or monolinguals, and indeed may be more heterogenous due to the greater complexity of a multilingual system.[1] Existing research on individual differences in cognitive ability has provided strong evidence for the role of language learning aptitude in the attainment of L2 proficiency in general and morphosyntactic knowledge in particular. In addition, there is evidence for a role of WM capacity, though findings are more mixed. By contrast, the potential influence of cognitive style remains under-researched. Furthermore, existing research has demonstrated that the impact of individual learner differences may vary with the instructional approach that is used, learners' level of proficiency, and the relative learning difficulty of the target structure. These findings are based on studies with L2 learners and/or studies which did not distinguish between L2 and L3/Lx learners, but there are no existing studies which have specifically worked with multilingual learners. In view of this state of current research, we formulated the following research questions for the present study:

1. Do multilingual learners benefit in equal measure from two types of explicit instruction (traditional/familiar vs. concept-based/unfamiliar) on the Spanish past tense/aspect distinction?
2. Do individual differences in language learning aptitude, working memory capacity, and cognitive style correlate with gain scores in the two treatment groups?
3. Does language learning experience relate to learners' performance?

Materials and methods

In order to address the research questions, we conducted a quasi-experimental study with multilingual adults. This section provides information about the participating learners, presents the targeted linguistic structure, the instructional conditions and experimental treatment, as well as the instruments used to assess the variables of interest. It also outlines the approach we took to analysing the data.

Participants

A total of 41 university-level learners (32 females, 9 males) enrolled in Spanish courses at a British University participated in the study. The participants were aged between 18 and 44 years (mean = 21) and came from 18 different L1 backgrounds: English (13 participants), Romanian (7), Italian, Lithuanian, Polish (3 each), Hungarian (2), Norwegian, French, Portuguese, Greek, Estonian, Latvian, Slovak, Russian, Czech, Bulgarian, Vietnamese and Cantonese (1 each). The participants had learned up to 6 other languages apart from Spanish (mean = 2.5).

Participants' language learning experience was gauged by means of a background questionnaire which consisted of questions about their current BA programme, year of study, years of Spanish learning at school and university, the language modules they were currently enrolled in (with the module codes revealing the language proficiency level), time spent in Spanish-speaking countries, other languages learned in a formal setting, i.e. at school or at university, and years of learning of each of these other languages. Participants achieved a mean score of 78% on the gap-fill pre-test used in the present study (please see below for details), indicating a solid level of knowledge. Proficiency in languages other than Spanish was not ascertained. Thus, we cannot state that all our participants were prototypical multilinguals according to the theoretical conceptualisation outlined in the background section above; regular and/or limiting cases are likely to be represented in our sample as well.

The participants were assigned to one of the two instructional conditions according to a matched ranking by language learning aptitude, WM capacity and cognitive style. As a consequence, the traditional instruction (TI) group ($n = 20$) and the concept-based instruction (CBI) group ($n = 21$) did not differ statistically in terms of aptitude, WM, or verbal/imagery orientation (all p-values > .29).

Target structure

The experimental treatment focused on teaching past tense/aspect marking in Spanish, which involves the distinction between the preterit and the imperfect. This structure is known to be difficult even for advanced learners (Negueruela, 2008; Salaberry, 2008). In order to use the preterit and imperfect accurately and appropriately, learners must appreciate the relationship between the linguistic form and the lexical information inherent in the semantic value of the verb. Moreover, contextual information can alter aspectual interpretation, so use of the preterit or the imperfect is not a categorical choice. In certain circumstances, either structure is possible, depending on the exact meaning a speaker wishes to convey. Thus, the contrastive use of preterit vs. imperfect goes beyond the core aspectual notions of 'completeness' vs. 'in progress' (Dominguez et al., 2013). In sum, learners must make the

appropriate formal choice for expressing the specific meaning they wish to communicate in a given context.

Instructional conditions and experimental treatment

We compared two instructional approaches labelled TI and CBI, with the latter representing an approach that was unfamiliar to participants and thus potentially challenging. The rationale behind CBI lies in Gal'perin's (1969) application of Vygotskian thought to pedagogy and involves three general principles. The first two principles state that concepts, or categories of meaning, should be seen as the minimal pedagogical unit and should be materialised by using, for instance, diagrams or charts. The full Gal'perian cycle involves a third principle that requires students to verbalise the target concepts (for an in-depth exploration of CBI, see Arievitch & Haenen, 2005; Lantolf & Poehner, 2014).

Due to the fact that Gal'perin's approach is not prescriptive, early research into CBI for L2 learning explored issues relating to the operationalisation of the model as such with a focus on how to materialise concepts (e.g., Lapkin et al., 2008; Negueruela, 2003, 2008; Swain et al., 2009). Furthermore, research into the potential effectiveness of the approach has reported its beneficial impact on L2 learning, although the studies in question tended to be descriptive, primarily documenting development through qualitative case studies (e.g., Gánem-Gutiérrez, 2016; Poehner & Infante, 2017). In some cases, descriptive statistics are reported based on pre-test/post-test designs, but only for a CBI group (e.g., Gánem-Gutiérrez & Harun, 2011; Kuepper & Feryok, 2020; Lantolf & Tsai, 2018). Comparative studies employing inferential statistics are still scarce, though are beginning to appear (e.g., Negrete Cetina, 2019; van Compernolle, 2018), and the present study follows this approach.

Both CBI and TI as operationalised in the present study were explicit instructional approaches in the sense that they made use of metalinguistic explanations and encouraged learners to construct and employ explicit knowledge about the target structure (Goo et al., 2015; Norris & Ortega, 2001). In essence, TI was similar to the instructional approach learners were familiar with from their current Spanish studies and previous or concurrent classroom instruction in other languages, while CBI was an approach that was new to them. The key characteristics of the two approaches are summarised in Table 1.

Table 2 shows examples of the metalinguistic explanations learners in the TI condition encountered, while Figure 1 exemplifies the kind of materials learners in the CBI condition were exposed to.

Table 1. Traditional instruction vs. concept-based instruction.

	Traditional instruction (TI)	Concept-based instruction (CBI)
Metalinguistic explanations	– Derived from (structural) pedagogical grammars – Emphasis on prescription, i.e. the use of rules is highlighted	– Derived from cognitive linguistics – Emphasis on the semantic implications of morphosyntactic choices
Use of explicit knowledge	– Step-by-step presentation of discrete pedagogical grammar rules – Rules take the form of verbal propositions, e.g. 'The preterit is used to refer to events in a narrative in the past.'	– Presentation of a linguistic concept (such as tense/aspect) in its entirety rather than incrementally – Linguistic concepts are presented by means of diagrams, pictures and charts

Table 2. Metalinguistic explanations from the TI condition.

Preterit	Imperfect
To refer to a single, complete action in the past	To refer to an ongoing action or state in the past with an unspecified time frame
To locate an event in the past	To refer to habitual actions in the past
To refer to a period of time that is considered finished	To describe people, places and objects in the past
To refer to the events in a narrative in the past	To set the scene in a narrative in the past (the actions would normally be in the preterit)
To refer to a past event completed before a second one	To express politeness

Note: Content for the TI condition was adapted from Butt and Benjamin (2000) and Turk and Zollo (2000).

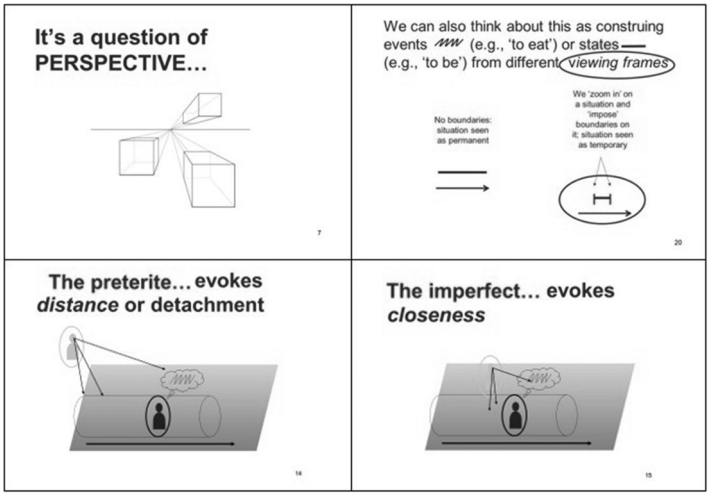

Figure 1. Example tutorial materials from the CBI condition.
Note: Content for the CBI condition was adapted from Castañeda Castro (2004, 2006), Doiz-Bienzobas (2002) and Radden & Dirven (2007).

As illustrated in the examples, the instructional materials differed in format and source of content, that is, discrete pedagogical grammar rules for TI versus explanations and diagrams derived from cognitive linguistics for CBI. TI content was characterised by a focus on form followed by practice that focused on meaning, while CBI content emphasised the role of form as a tool to convey nuances in meaning, i.e. form and meaning were always presented in conjunction. Otherwise, the tasks used were comparable in that they were presented visually, employed discrete multiple-choice items, relied on a closed response format and incorporated implicit feedback, as described below. The administration procedures and time on task were also the same in both conditions.

Specifically, both groups completed two 90-minute computer-based treatment sessions on separate days. The sessions comprised a tutorial plus a practice task (Day 1) and revision plus a practice task (Day 2). All materials for the treatment sessions were designed and administered using MS Powerpoint. Participants worked individually at computer stations under supervised conditions. In the tutorial stage, they worked their way through a series of slides in 'show' mode. The tutorial was followed by a practice task consisting of a series of pictures illustrating a story. Participants had to click on the button (conjugated verbs in the TI condition, diagrams in the CBI condition) which would best describe what was depicted in the slides, as illustrated in Figure 2. Feedback was implicit, i.e. if the correct choice was made, the next slide would appear, whereas an incorrect choice would take the participant

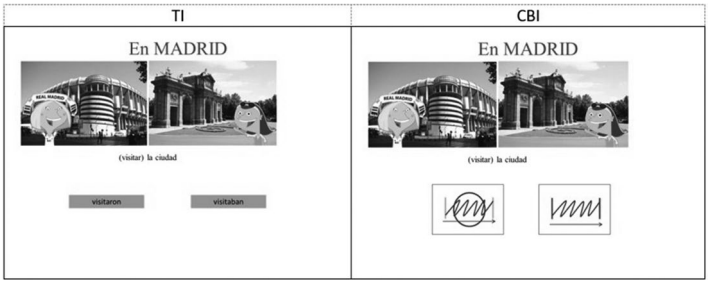

Figure 2. Example practice materials.
Note: Materials for the practice task were adapted from the Spanish Learner Language Oral Corpora (SPLLOC) project (http://www.splloc.soton.ac.uk) and Castañeda Castro (2006).

back to the beginning of the task in order for them to try again. The revision session on Day 2 was based on the tutorial materials from Day 1, followed by a re-run of the practice task accomplished on Day 1.

Instruments

In order to assess participants' knowledge of the target structure, two matched tests prior to and following the instructional treatment were administered: a gap-fill test and a test of metalinguistic knowledge. The gap-fill test allowed participants to draw on any type of knowledge at their disposal, i.e., implicit knowledge, explicit knowledge or both in conjunction. The test of metalinguistic knowledge was aimed at eliciting explicit knowledge about past tense/aspect marking in Spanish. In the present study, explicit knowledge was defined as knowledge that a learner is consciously aware of and that can potentially be articulated in a verbal statement (Ellis, 2004; Hulstijn, 2005; Roehr-Brackin, 2018). Explicit knowledge is represented declaratively, accessed via controlled processing (Hulstijn, 2005) and can be called up on demand (Dörnyei, 2009). By contrast, implicit knowledge is understood as intuitive knowledge that is accessed via automatic processing, can be used in performance, but cannot be brought into awareness or be articulated (Dörnyei, 2009; Hulstijn, 2005). We may thus also describe implicit knowledge as knowledge that is not explicit.

The gap-fill test was presented as a story which comprised 16 target items and 6 distractors at sentence level. Participants had to insert a verb in either its preterit or its imperfect form into each gap, as appropriate to the communicative context. The infinitive form of the respective verbs was provided. The test was scored dichotomously, resulting in a maximum possible score of 16. The distractors were not scored. The gap-fill test showed moderate but arguably still reasonable reliability, given the relatively small number of items: Cronbach's alpha = .62 (pre-test) and = .68 (post-test).

The test of metalinguistic knowledge consisted of five parts with 39 tasks in total. The test was aimed at assessing participants' explicit understanding of terminology as well as relationships between form and meaning pertaining to the target structure. Parts A and D elicited knowledge of terminology and concepts related to the preterit and imperfect through an open question format, while Part E asked learners to match the target forms with appropriate conceptual categories (e.g., the concept of boundedness with the preterit). Part B required sentence-level correction and metalinguistic explanation, thus targeting knowledge of pedagogical grammar rules. Part C assessed participants' awareness of semantic contrasts evoked by the two forms. Table 3 provides examples from each test part. The

Table 3. Sample items from the metalinguistic knowledge test.

Part	No. of tasks	Maximum possible score	Example
A	1	2	When talking about past events in Spanish we use two contrasting verb forms. Please name them.
B	8	40	*Sonaba el teléfono cuando estaba leyendo.* Correction: Explanation:
C	3	9	*Cuando **volvíamos**a cas a, nos encontramos a Juan.* *Cuando **volvimos**a cas a, nos encontramos a Juan.* How would you explain the difference in meaning between these two sentences?
D	5	15	Can you explain what 'aspect' is?
E	20	20	Boundaries of situation in focus (Targeted answer: Preterit)

maximum possible score for the test of metalinguistic knowledge was 86. The test showed reasonable reliability: Cronbach's alpha = .78 (pre-test) and = .82 (post-test).

After the experimental treatment, the participants completed a short questionnaire aimed at gauging their attitudes towards the instructional approach they had experienced. Items focused on whether learners considered the instruction helpful, whether they found the metalinguistic explanations easy to understand, whether the instructional materials had helped them better understand the distinction between the preterit and the imperfect, and whether they preferred the instructional materials in their regular textbook. The items were presented as statements and answers were provided on a five-point Likert scale ranging from 'strongly agree' to 'strongly disagree'.

The individual difference variables measured in the present study were language learning aptitude, WM capacity, and cognitive style on the verbal/imagery dimension. Participants were tested on these variables before the experimental treatment commenced.

Cognitive style refers to "an individual's preferred and habitual approach to both organizing and representing information" (Riding, 2001, p. 48). Riding's model of cognitive style comprises two continua: wholist/analytic and verbal/imagery. We focused on the latter because the materials used in the two instructional conditions differed according to whether they relied on primarily verbal propositions (TI) or primarily graphic, diagrammatic depictions (CBI). Overall, individuals with a verbal preference tend to represent information during thinking verbally, while individuals with an imagery preference tend to think in mental images. It could therefore be hypothesised that learners with a verbal preference might do better in and/or express more positive attitudes towards the TI condition, whereas learners with an imagery preference might do better in and/or express more positive attitudes towards the CBI condition. We assessed cognitive style through the Verbal-Imagery Cognitive Style (VICS) measure (Peterson et al., 2005a, 2005b), which is a computer-based test measuring preferences on each of the two dimensions by means of median response times to a range of stimuli presented either in verbal format (e.g., the word 'nail') or in pictorial format (e.g., the image of a nail), thus favouring either verbal or imagery processing. Response times are logged automatically, and the measure takes about 15 minutes to complete. Scores take the form of an automatically calculated V/I ratio. A ratio closer to 0 indicates a verbal preference, a ratio closer to (or above) 2 indicates an imagery preference, while ratios between .8 and 1 suggest no preference in either direction (Peterson, 2003).

Language learning aptitude was assessed by means of the LLAMA aptitude test battery (Meara, 2005; Rogers et al., 2016). This computer-administered measure comprises four

subtests: LLAMA-B – Vocabulary Learning tests the ability to learn new words; LLAMA D – Sound Recognition tests the ability to recognise patterns in spoken language; LLAMA E – Sound-Symbol Association tests the ability to form new sound-symbol associations, i.e. phonetic coding ability in Carollian terms; and LLAMA F – Grammatical Inferencing tests the ability to induce the rules of an unknown language. The test is scored automatically and takes about 30 minutes to complete. The maximum possible score is 375.

A backward digit span (BDS) test was used to assess participants' complex WM capacity. The measure requires test takers to listen to sets of sequences of digits which they must then repeat in reverse order at the end of each set. The sequences increase in length as the test progresses, starting with 3 digits (e.g., 3 – 7 – 1), then 4 digits, etc. up to a maximum of 9 digits. The test thus gets increasingly demanding. As the measure requires test takers to not only recall digits, but produce them in reverse order, it draws on both storage and processing of information (Juffs & Harrington, 2011; Sáfár & Kormos, 2008). The BDS test is typically administered in the test taker's L1 to avoid any confound with L2 proficiency. As the participants in our study spoke a variety of L1s, the sequences of digits were shown in visual format by means of an automatically timed Powerpoint presentation. Stimuli appeared at a speed of one digit per second. Each set contained six sequences of digits. Participants were instructed to repeat the digits in reverse order in their L1; their oral responses were audio-recorded and scored afterwards. Participants were awarded 3 points if they correctly repeated in reverse order at least two of the first three sequences in the first set and thereafter 0.5 points for each subsequent three sequences in a set, if at least two out of three were correctly repeated in reverse order. This resulted in a maximum possible score of 9.5.

Data analysis

Normality of distribution of scores on all variables was assessed by means of one-sample KS tests. As a number of variables differed significantly from a normal distribution, analyses involving these variables relied on non-parametric tests (Spearman correlations, Mann Whitney U tests for group comparisons, Wilcoxon Signed Rank tests for pre-test/post-test comparisons). In the case of variables with a normal distribution, standard parametric tests were used (Pearson correlations, independent *t*-tests for group comparisons, paired samples *t*-tests for pre-test/post-test comparisons). The alpha level was set at .05, with Bonferroni corrections applied for multiple comparisons.

Table 4 shows the descriptive statistics for the LLAMA, VICS and BDS. VICS results are displayed in terms of V/I ratio. Recall that values closer to 0 indicate a verbal preference and values closer to 2 or above indicate an imagery preference. As the participants had been assigned to instructional conditions based on a matched ranking, the two groups show a similar performance on the individual difference measures. With regard to language learning aptitude, LLAMA D was the most difficult subtest, while LLAMA E was the easiest. This is in keeping with existing work (Rogers et al., 2016) and thus unsurprising. The V/I ratio range indicates that there were learners with both verbal and imagery preferences in both treatment groups, as intended, though the mean is clearly in the neutral spectrum. There was also a good spread of LLAMA total scores and WM capacity in both groups, with backward digit spans ranging from 3.0 to 7.0.

Table 4. Descriptive statistics for the individual difference measures.

	TI Group ($n = 20$)							CBI Group ($n = 21$)						
	LLAMA total	LLAMA B	LLAMA E	LLAMA F	LLAMA D	V/I Ratio	BDS	LLAMA total	LLAMA B	LLAMA E	LLAMA F	LLAMA D	V/I Ratio	BDS
Mean	241.75	56.25	88.50	63.50	35.25	.94825	4.711	239.50	55.95	90.95	57.50	38.81	.89424	4.238
Median	240.00	55.00	95.00	70.00	35.00	.93650	4.500	242.50	55.00	90.00	60.00	40.00	.85100	4.000
SD	42.466	11.571	13.089	22.070	15.768	.169525	1.0317	42.609	21.072	10.911	16.819	15.961	.154114	.8459
Minimum	150	40	70	20	5	.714	3.5	145	20	70	20	5	.598	3.0
Maximum	320	75	100	100	60	1.366	7.0	305	90	100	80	70	1.183	6.0
Max. possible	375	100	100	100	75	n/a	9.5	375	100	100	100	75	n/a	9.5
Mean %	65	56	89	64	47	n/a	50	64	56	91	58	39	n/a	45

Note: V/I = Verbal/imagery, BDS = Backward digit span.

Table 5. Descriptive statistics for the language tests.

	TI Group ($n = 20$)						CBI Group ($n = 21$)					
	Gap-fill pre-test	Gap-fill post-test	Gap-fill gains	MLK pre-test	MLK post-test	MLK gains	Gap-fill pre-test	Gap-fill post-test	Gap-fill gains	MLK pre-test	MLK post-test	MLK gains
Mean	12.05	10.70	−1.35	33.10	51.00	17.90	12.86	12.00	−.86	35.24	55.52	20.24
Median	12.00	11.00	−1.00	32.00	51.50	15.00	13.00	12.00	−1.00	33.00	55.00	20.00
SD	2.724	3.097	1.785	8.284	9.830	8.955	2.128	2.846	2.032	8.671	9.516	7.321
Minimum	5	5	−6	22	34	4	7	7	−5	18	30	9
Maximum	16	16	1	49	70	38	16	16	2	47	69	38
Max. possible	16	16	16	86	86	86	16	16	16	86	86	86
Mean %	75	69	−6	39	59	20	80	75	−5	41	65	23

Note: MLK = Metalinguistic knowledge.

Results

The first research question asked whether multilingual participants would benefit in equal measure from two types of explicit instruction (traditional vs. concept-based) on the Spanish past tense/aspect distinction. We began to address this question by scrutinising the descriptive statistics for the two instructional groups on the gap-fill test and the test of metalinguistic knowledge assessing learners' knowledge of the Spanish past tense/aspect distinction, as shown in Table 5.

The descriptive statistics suggest that the two groups performed similarly on the two measures. The fact that the gap-fill test resulted in negative gain scores is surprising, although this may be due to the high scores already obtained at pre-test. The test of metalinguistic knowledge shows the expected pattern of positive gains of about 20% between pre- and post-test. Inferential statistics (Bonferroni-corrected) confirm that the TI group made significant losses between pre-test and post-test on the gap-fill test, $t(19) = 3.382$, $p = .003$, and also significant gains between pre-test and post-test on the MLK test, $t(19) = -8.939$, $p < .001$. Conversely, the CBI group's losses on the gap-fill test between pre-test and post-test were not statistically significant, although a trend is in evidence, $z = -1.712$, $p = .087$. The CBI group also made significant gains on the MLK test between pre-test and post-test, $t(20) = -12.649$, $p < .001$. A comparison between groups shows no statistical differences in terms of (negative) gain scores on the gap-fill test, $t(39) = -.823$, $p = .598$, or in terms of (positive) gain scores on the MLK test, $t(39) = -.917$, $p = .409$.

The second research question asked whether individual differences in language learning aptitude, working memory capacity or cognitive style would correlate with gain scores in the two treatment groups. Correlational analyses (Bonferroni-corrected) for the TI group revealed a trend towards a negative correlation between gain scores on the gap-fill test and V/I ratio (*rho* $= -.50$, $p = .025$), indicating that participants leaning towards a verbal stylistic preference showed a trend towards making greater gains, while participants with an imagery stylistic preference showed a trend towards fewer gains. There were no other statistical relationships with individual difference variables in this group (all coefficients < .32, all *p*-values > .41). Correlational analyses (Bonferroni-corrected) for the CBI group yielded no statistical relationships or trends between any gain scores and any of the individual difference variables ($r = -.31$, $p = .19$ for LLAMA total and metalinguistic gains, all other coefficients < .14, all other *p*-values > .70).

Participants' responses to the post-treatment questionnaire showed positive attitudes overall to the instructional approaches the learners had experienced, with mean scores at 4.34 (TI group) and 4.24 (CBI group). A maximum of 5 indicates the most positive attitude possible. There were no statistical differences between the two groups on any of the attitude questions (all *p*-values > .36).

The third research question asked whether language learning experience would relate to learners' performance. Correlational analyses performed for the sample as a whole yielded no statistical relationship between the number of languages apart from Spanish that participants had learned and gains on the gap-fill test, the metalinguistic knowledge test, or performance on the measures of aptitude, WM or verbal/imagery preference (*rho* = .22, $p = .17$ for LLAMA total, all other coefficients < .13, all other *p*-values > .40). In addition, there was no statistical difference in terms of gain scores on the gap-fill test, $t(39) = .386$, $p = .701$, or the metalinguistic knowledge test, $t(39) = .790$, $p = .434$, between learners with a Romance L1 ($n = 12$) and learners with a non-Romance L1 ($n = 31$).

Discussion and conclusions

In summary, our study with multilingual adult learners who were exposed to two different explicit instructional approaches focusing on the Spanish past tense/aspect distinction revealed that both groups made equal significant gains in terms of metalinguistic knowledge. In other words, TI and CBI were equally effective instructional methods for the purpose of building explicit knowledge about the target structure. Somewhat surprisingly, both groups suffered equal losses on the gap-fill test, which was aimed at assessing learners' use of the target structure, with the loss suffered by the TI group reaching statistical significance and the loss suffered by the CBI group showing as a trend.

The groups exhibited no differences in attitudes towards the instructional method they had experienced, with positive views being the norm. Participants performed equally regardless of whether they had a Romance or a non-Romance L1, and the number of languages known apart from Spanish was not associated with performance on any of the linguistic or individual difference measures.

Individual differences in language learning aptitude, complex WM and cognitive style did not correlate with gain scores, although the TI group showed a trend towards an association between a more verbal stylistic orientation and greater gains on the gap-fill test, or, formulated alternatively, a more imagery stylistic orientation and fewer gains on the gap-fill test.

Taking this last point first, the observed trend is in line with expectations, since the explicit explanations of the target structure in the TI condition relied on verbal propositions. It is worth noting, however, that the trend was observed for gains on the gap-fill test, not for gains on the metalinguistic knowledge test where explicit information relying on verbal propositions would be expected to be applied first and foremost. It is not immediately obvious why no association with stylistic orientation was found in this instance. A possible explanation might be that remembering explicit knowledge taught during the instructional treatment did not benefit from a verbal orientation, but bringing to bear this knowledge to aid in the performance on the gap-fill test was more demanding and thus put participants with a matching stylistic preference at an advantage. However, this must remain speculative, since we do not know whether and to what extent specific learners drew on explicit knowledge when completing the gap-fill test. Moreover, it is noteworthy that there was no corresponding trend in the CBI group. We might have expected an association between an imagery preference and greater gains on either or both of the language measures, but this was not found. A possible interpretation would be to suggest that the materials in the CBI treatment did not require a matching stylistic orientation for being processed successfully, whereas the materials in the TI treatment tended to do so to a greater extent.

Contrary to expectation, we observed negative gains on the gap-fill test in both groups, significantly so in the TI group and marginally in the CBI group. At first glance, this result is counter-intuitive, especially as the participants made significant positive gains in terms of metalinguistic knowledge. A possible explanation may be found in the high pre-test scores participants achieved on the gap-fill test. They already performed well before the experimental treatment began, either because they had a good (implicit?) grasp of the target structure, or because the test was perhaps a little too easy for them, given its closed-ended and highly controlled format, or indeed because both of these conditions applied and reinforced each other.

Participants' post-test performance suggests that the experimental treatment may have triggered a restructuring in learners' knowledge systems. Their metalinguistic knowledge increased significantly, and their 'shaky' gap-fill test performance at post-test may reflect the beginnings of a knock-on effect on their linguistic knowledge. According to complexity and dynamic systems theory which emphasises the non-linear nature of language learning, variability in performance observed at a particular point in time is an indication of change and often a precursor of development (Larsen-Freeman & Cameron, 2008; Verspoor & Behrens, 2011). In this sense, variability, and indeed a (temporary) drop in accuracy, may be interpreted as a harbinger of progress to come (de Bot et al., 2007; de Bot & Larsen-Freeman, 2011).

It is also noteworthy that the gap-fill test showed greater reliability at post-test, suggesting that less guessing took place than at pre-test. Overall, participants may have begun to question and revise previously assumed knowledge about the target structure in light of the explicit information gathered during the instructional treatment. Some of the questioning and revision led to incorrect responses, to be sure, but nonetheless this can be indicative of a reorganisation of existing knowledge representations. The observed drop in performance is clearly only part of the picture and may thus be temporary only. While this interpretation relies on circumstantial inferencing, the observed significant increase in metalinguistic knowledge provides some supporting evidence for such a viewpoint.

We did not find any significant associations between individual differences in cognitive ability, operationalised via measures of language learning aptitude and WM capacity, and gains on either of the language measures in either instructional condition. Scores on all measures showed a good range, thus confirming that our sample was sufficiently heterogeneous for significant relationships to emerge in principle. Moreover, given that the target structure was difficult and that one of the instructional conditions (CBI) was unfamiliar and thus potentially challenging for the participants, one might have expected aptitude and/or WM to have an impact – yet this was not the case. Possible reasons may be found in the other factors that are known to play a role in the complex interaction of variables that contribute to multilingual language learning and use.

First, the participating learners were experienced multilinguals most of whom had learned at least one other language apart from Spanish and some several more; they had been exposed to different instructional settings, and they had thus accumulated considerable language learning experience. This may have resulted in a levelling effect (Erlam, 2005; Serafini & Sanz, 2016), which cancelled out any influence of relatively higher as opposed to relatively lower cognitive abilities. Participants' communicative and strategic skills would have been honed to a considerable extent, and the M-factor (Cenoz & Jessner, 2009; Jessner, 2008, 2014) was likely in play. Once such a level of experience has been reached, it may simply not matter anymore whether a new instructional method is encountered. It was an explicit approach after all, and multilinguals' metalinguistic abilities are known to be well-developed. Individual differences in aptitude and WM may have lost their relevance at this advanced stage.

Second, learners' relatively high level of knowledge of the target structure and their relatively high level of Spanish proficiency overall, as evidenced by their gap-fill test performance at pre-test, may be another reason for the attenuation of any potential effect of cognitive individual differences. As previous research has shown, aptitude appears to be most relevant at lower proficiency levels (Li, 2015, 2016; Rodríguez Silva & Roehr-Brackin, 2016), and so does WM (Juffs & Harrington, 2011; Rodríguez Silva & Roehr-Brackin, 2016; Serafini & Sanz, 2016).

Third, studies comparing different instructional approaches which identified any predictive power of aptitude and/or WM tended to find such effects on oral rather than written outcome measures (Rodríguez Silva & Roehr-Brackin, 2016; Tomak & Roehr-Brackin, 2017). This suggests that higher levels of cognitive ability convey advantages on speaking tasks, which by their very nature require fast access to any existing knowledge representations. By contrast, all learners with a certain level of instructional experience may be able to handle written measures that are less time-pressured and thus allow for explicit knowledge to be accessed. Under such conditions, a higher level of aptitude and/or WM may no longer convey any additional benefits because these abilities are not sufficiently taxed by written tasks, especially if a certain proficiency level has been reached, and even more so if the learners in question are seasoned multilinguals with considerable exposure to different languages under their belts, as in the case of the present study.

Limitations and future research

In sum, the findings reported here have both substantiated and extended existing work in that we have identified an apparent levelling effect of language learning experience in instructed multilinguals who were tested by means of written outcome measures. In our participants, the M-factor was seemingly sufficiently powerful to compensate for an unfamiliar instructional approach (CBI) targeting a challenging linguistic feature (Spanish past tense/aspect), resulting in similar performance compared with learners experiencing a familiar instructional approach (TI).

At the same time, we must acknowledge the limitations of our study. The number of participants (N = 41) was relatively small, which meant that we could not include a control group. Moreover, the instructional treatment was relatively short (2 × 90 minutes), although this is not unusual in laboratory settings using computer-administered instruction. Participants' strong performance on the gap-fill task at pre-test suggests that a more challenging measure would have been desirable, with an oral test requiring online processing probably most appropriate to bring out the potential role of individual differences in cognitive abilities. Therefore, future research investigating the interplay of individual learner factors and different instructional approaches in experienced language learners would benefit from using a range of outcome measures which include time-pressured tests that do not allow so readily for the use of explicit knowledge. Moreover, the inclusion of learners at different levels of proficiency, including learners at lower levels, would be desirable, so the threshold of any levelling effect which seemingly neutralises individual differences in cognitive ability can be identified.

Note

1. As pointed out by an anonymous reviewer, it may also be the case that greater metalinguistic awareness may attenuate individual differences between multilinguals, similarly to the attenuating effect of increased proficiency.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Karen Roehr-Brackin is a Reader in the Department of Language and Linguistics at the University of Essex. Her research focuses on additional language learning in children and adults. Her specific interests are explicit and implicit knowledge and learning in the classroom as well as the role of individual learner differences such as language learning aptitude and working memory.

Gabriela Adela Gánem-Gutiérrez is a Senior Lecturer in applied linguistics in the Department of Language and Linguistics at the University of Essex. Her main research interests include L2 pedagogy, explicit L2 learning, sociocultural theory, applied cognitive linguistics for language teaching, and CALL.

Lexa Olivera-Smith is a Senior Lecturer in Spanish in the Department of Language and Linguistics at the University of Essex. She teaches higher intermediate Spanish and a range of translation, interpreting and subtitling modules at both undergraduate and postgraduate level. She is the course director of the MA Translation & Professional Practice and MA Translation & Literature. Her main research interests are audiovisual translation, literary translation and pedagogy of translation.

María Teresa Torres-Marín is a Senior Lecturer in Spanish in the Department of Language and Linguistics at the University of Essex. She teaches Spanish language from ab initio to advanced level. Her main interests include L2 pedagogy, Spanish cinema and the use of short films in language teaching.

ORCID

Karen Roehr-Brackin http://orcid.org/0000-0002-1152-5852

References

Arievitch, I. M., & Haenen, J. P. P. (2005). Connecting sociocultural theory and educational practice: Galperin's approach. *Educational Psychologist*, *40*(3), 155–165. https://doi.org/10.1207/s15326985ep4003_2

Baker, C. (2006). *Foundations of bilingual education and bilingualism* (4th ed.). Multilingual Matters.

Berthele, R. (2021). The extraordinary ordinary: Re-engineering multilingualism as a natural category. *Language Learning*, *71*(S1), 80–120. https://doi.org/10.1111/lang.12407

Bialystok, E. (2001). *Bilingualism in development: Language, literacy, and cognition*. Cambridge University Press.

Biedroń, A., & Szczepaniak, A. (2012). Working memory and short-term memory abilities in accomplished multilinguals. *The Modern Language Journal*, *96*(2), 290–306. https://doi.org/10.1111/j.1540-4781.2012.01332.x

Butt, J., & Benjamin, C. (2000). *A new reference grammar of modern spanish* (3rd ed). London: Arnold.

Castañeda Castro, A. (2004). Una visión cognitiva del sistema temporal y modal del verbo en español. *ELUA. Estudios de Lingüística Universidad de Alicante*, (Anexo 2), *5*, 55–71. https://doi.org/10.14198/ELUA2004.Anexo2.04

Castañeda Castro, A. (2006). Aspecto, perspectiva y tiempo de procesamiento en la oposición imperfecto/indefinido del español: Ventajas explicativas y aplicaciones pedagógicas. *RAEL: Revista Electrónica de Lingüística Aplicada*, *5*, 107–140.

Cenoz, J., & Jessner, U. (2009). The study of multilingualism in educational contexts. In L. Aronin & B. Hufeisen (Eds.), *The exploration of multilingualism* (pp. 121–138). John Benjamins.

de Bot, K., & Larsen-Freeman, D. (2011). Researching second language development from a dynamic systems theory perspective. In M. H. Verspoor, K. de Bot, & W. Lowie (Eds.), *A dynamic approach to second language development: Methods and techniques* (pp. 5–23). John Benjamins.

de Bot, K., Lowie, W., & Verspoor, M. (2007). A dynamic systems theory approach to second language acquisition. *Bilingualism: Language and Cognition*, *10*(01), 7–21. https://doi.org/10.1017/S1366728906002732

Doiz-Bienzobas, A. (2002). The preterit and the imperfect as grounding predications. In B. Frank (Ed.), *Grounding: The epistemic footing of deixis and reference* (pp. 299–348). Mouton de Gruyter.

Dominguez, L., Tracy-Ventura, N., Arche, M. J., Mitchell, R., & Myles, F. (2013). The role of dynamic contrasts in the L2 acquisition of Spanish past tense morphology. *Bilingualism: Language and Cognition*, *16*(3), 558–577. https://doi.org/10.1017/S1366728912000363

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Erlbaum.

Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford University Press.

Ellis, R. (2004). The definition and measurement of L2 explicit knowledge. *Language Learning*, *54*(2), 227–275. https://doi.org/10.1111/j.1467-9922.2004.00255.x

Erlam, R. (2005). Language aptitude and its relationship to instructional effectiveness in second language acquisition. *Language Teaching Research*, *9*(2), 147–171. https://doi.org/10.1191/1362168805lr161oa

Gal'perin, P. I. (1969). Stages in the development of mental acts. In M. Cole & I. Maltzman (Eds.), *A handbook of contemporary Soviet psychology* (pp. 249–273). Basic Books.

Gánem-Gutiérrez, G. A. (2016). Enhancing metalinguistic knowledge: Preterite and imperfect in L2 Spanish. *Language and Sociocultural Theory*, *3*(1), 27–54. https://doi.org/10.1558/lst.v3i1.28803

Gánem-Gutiérrez, G. A., & Harun, H. (2011). Verbalisation as a mediational tool for understanding tense-aspect marking in English: An application of Concept-Based Instruction. *Language Awareness*, *20*(2), 99–119. https://doi.org/10.1080/09658416.2010.551125

Gombert, J. E. (1992). *Metalinguistic development*. Hemel Hempstead: Harvester.

Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning: Norris & Ortega (2000) revisited and updated. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 443–482). John Benjamins.

Hufeisen, B. (2018). Models of multilingual competence. In A. Bonnet & P. Siemund (Eds.), *Foreign language education in multilingual classrooms* (pp. 173–189). John Benjamins.

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition*, *27*(02), 129–140. https://doi.org/10.1017/S0272263105050084

Jessner, U. (1999). Metalinguistic awareness in multilinguals: Cognitive aspects of third language learning. *Language Awareness*, *8*(3–4), 201–209. https://doi.org/10.1080/09658419908667129

Jessner, U. (2006). *Linguistic awareness in multilinguals: English as a third language*. Edinburgh University Press.

Jessner, U. (2008). A DST model of multilingualism and the role of metalinguistic awareness. *The Modern Language Journal*, *92*(2), 270–283. https://doi.org/10.1111/j.1540-4781.2008.00718.x

Jessner, U. (2014). On multilingual awareness or why the multilingual learner is a specific language learner. In M. Pawlak & L. Aronin (Eds.), *Essential topics in applied linguistics and multilingualism: Studies in honor of David Singleton* (pp. 175–184). Springer.

Juffs, A., & Harrington, M. (2011). Aspects of working memory in L2 learning. *Language Teaching*, *44*(2), 137–166. https://doi.org/10.1017/S0261444810000509

Kuepper, M. C., & Feryok, A. (2020). Concept-based pragmatics instruction: Teaching German address pronouns to New Zealand tertiary students. *Language and Sociocultural Theory*, *6*(2), 158–183. https://doi.org/10.1558/lst.37359

Lantolf, J. P., & Poehner, M. E. (2014). *Sociocultural theory and the pedagogical imperative in L2 education: Vygotskian praxis and the research/practice divide*. Routledge.

Lantolf, J. P., & Tsai, M. H. (2018). L2 developmental education and systemic theoretical instruction: The case of English verb + noun collocations. In A. E. Tyler, L. Ortega, M. Uno, & H. In Park (Eds.), *Usage-inspired L2 instruction* (pp. 29–53). John Benjamins.

Lapkin, S., Swain, M., & Knouzi, I. (2008). Postsecondary French as a second language students learn the grammatical concept of voice: Study design, materials development, and pilot data. In J. P.

Lantolf and M. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 228–255). Equinox.

Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford University Press.

Li, S. (2015). The associations between language aptitude and second language grammar acquisition: A meta-analytic review of five decades of research. *Applied Linguistics*, *36*(3), 385–408. https://doi.org/10.1093/applin/amu054

Li, S. (2016). The construct validity of language aptitude: A meta-analysis. *Studies in Second Language Acquisition*, *38*(4), 801–842. https://doi.org/10.1017/S027226311500042X

Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, *21*(4), 861–883. https://doi.org/10.3758/s13423-013-0565-2

Littlemore, J. (2001). An empirical study of the relationship between cognitive style and the use of communication strategy. *Applied Linguistics*, *22*(2), 241–265. https://doi.org/10.1093/applin/22.2.241

Meara, P. (2005). *Llama language aptitude tests*. University of Wales Swansea.

Negrete Cetina, M. (2019). *Enhancing EFL teacher trainees' cognition through Systemic Theoretical Instruction (STI)* [Unpublished PhD thesis]. University of Essex.

Negueruela, E. (2003). *A sociocultural approach to teaching and researching second language: Systemic-theoretical instruction and second language development* [Unpublished PhD thesis]. Pennsylvania State University.

Negueruela, E. (2008). Revolutionary pedagogies: Learning that leads (to) second language development. In J. P. Lantolf & M. E. Poehner (Eds.), *Sociocultural theory and the teaching of second languages* (pp. 189–227). Equinox.

Norris, J. M., & Ortega, L. (2001). Does type of instruction make a difference? Substantive findings from ameta-analyticreview.*LanguageLearning*,*51*(1),157–213.https://doi.org/10.1111/j.1467-1770.2001.tb00017.x

Ortega, L., Tyler, A. E., Park, H. I., & Uno, M. (Eds.). (2016). *The usage-based study of language and multilingualism*. Georgetown University Press.

Peterson, E. R. (2003). *Verbal imagery cognitive styles test & extended cognitive style analysis – wholistic analytic test: Administration guide*. University of Edinburgh.

Peterson, E. R., Deary, I. J., & Austin, E. J. (2005a). A new measure of verbal-imagery cognitive style: VICS. *Personality and Individual Differences*, *38*(6), 1269–1281. https://doi.org/10.1016/j.paid.2004.08.009

Peterson, E. R., Deary, I. J., & Austin, E. J. (2005b). Are intelligence and personality related to verbal-imagery and wholistic-analytic cognitive styles?*Personality and Individual Differences*, *39*(1), 201–213. https://doi.org/10.1016/j.paid.2005.01.009

Poarch, G. J., & van Hell, J. G. (2012). Executive functions and inhibitory control in multilingual children: Evidence from second-language learners, bilinguals, and trilinguals. *Journal of Experimental Child Psychology*, *113*(4), 535–551. https://doi.org/10.1016/j.jecp.2012.06.013

Poehner, M. E., & Infante, P. (2017). Mediated development: A Vygotskian approach to transforming second language learner abilities. *TESOL Quarterly*, *51*(2), 332–357. https://doi.org/10.1002/tesq.308

Radden, G., & Dirven, R. (2007). *Cognitive English grammar*. John Benjamins.

Riding, R. J. (2001). The nature and effects of cognitive style. In R. J. Sternberg & L.-F. Zhang (Eds.), *Perspectives on thinking, learning, and cognitive styles* (pp. 47–72). Erlbaum.

Rodríguez Silva, L. H., & Roehr-Brackin, K. (2016, June 10–11). Language learning aptitude and working memory as predictors of instructed adult L2 learners' explicit and implicit L2 knowledge. *Paper Presented at the Cognitive Approaches to Language Pedagogy 2 Conference*, Switzerland.

Roehr, K. (2008). Linguistic and metalinguistic categories in second language learning. *Cognitive Linguistics*, *19*(1), 67–106. https://doi.org/10.1515/COG.2008.005

Roehr-Brackin, K. (2018). *Metalinguistic awareness and second language acquisition*. Routledge.

Rogers, V., Meara, P., Aspinall, R., Fallon, L., Goss, T., Keey, E., & Thomas, R. (2016). Testing aptitude: Investigating Meara's (2005) LLAMA tests. *EUROSLA Yearbook*, *16*, 179–210. https://doi.org/10.1075/eurosla.16.07rog

Rosch, E. (1978). Principles of categorization. In E. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Erlbaum.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. https://doi.org/10.1016/0010-0285(75)90024-9

Sáfár, A., & Kormos, J. (2008). Revisiting problems with foreign language aptitude. *IRAL*, *46*, 113–136.

Salaberry, M. R. (2008). *Marking past tense in second language acquisition: A theoretical model*. Continuum.

Serafini, E. J., & Sanz, C. (2016). Evidence for the decreasing impact of cognitive ability on second language development as proficiency increases. *Studies in Second Language Acquisition*, *38*(4), 607–646. https://doi.org/10.1017/S0272263115000327

Sternberg, R. J., & Grigorenko, E. L. (1997). Are cognitive styles still in style? American Psychologist, *52*, 700–712.

Swain, M., Lapkin, S., Knouzi, I., Suzuki, W., & Brooks, L. (2009). Languaging: University students learn the grammatical concept of voice in French. *Modern Language Journal*, *93*(1), 6–30.

Tomak, I., & Roehr-Brackin, K. (2017, April). The role of language learning aptitude and type of instruction in the development of explicit and implicit knowledge of L2 English articles. Paper presented at the Thinking, Learning, Doing – Usage-Based Perspectives on L2 Learning (TDL3) Conference (pp. 19–21). Munich, Germany.

Turk, P., & Zollo, M. (2000). *Acción gramática* (2nd ed.). Hodder Education.

van Compernolle, R. A. (2018). Learning the sociopragmatics of an artificial language: Comparing rule-based and concept-based instruction. *Language and Sociocultural Theory*, *5*(1), 1–20. https://doi.org/10.1558/lst.33770

Verspoor, M., & Behrens, H. (2011). Dynamic systems theory and a usage-based approach to second language development. In M. H. Verspoor, K. de Bot, & W. Lowie (Eds.), *A dynamic approach to second language development: Methods and techniques* (pp. 25–38). John Benjamins.

Wen, Z. (2012). Working memory and second language learning. *International Journal of Applied Linguistics*, *22*(1), 1–22. https://doi.org/10.1111/j.1473-4192.2011.00290.x

Wen, Z., Biedroń, A., & Skehan, P. (2017). Foreign language aptitude theory: Yesterday, today and tomorrow. *Language Teaching*, *50*(1), 1–31. https://doi.org/10.1017/S0261444816000276

Wesche, M. B. (1981). Language aptitude measures in streaming, matching students with methods, and diagnosis of learning problems. In K. C. Diller (Ed.), *Individual differences and universals in language learning aptitude* (pp. 119–154). Newbury House.

Ziętek, A. A., & Roehr, K. (2011). Metalinguistic knowledge and cognitive style in Polish classroom learners of English. *System*, *39*(4), 417–426. https://doi.org/10.1016/j.system.2011.05.005