Are certain types of instruction better for certain learners?

Pavel Trofimovich*, Patsy M. Lightbown, Randall Halter

Centre for the Study of Learning and Performance, Concordia University, Montréal, Canada

Abstract

The goal of this study was to investigate a possible link between second language (L2) learners' background variables and the type of instruction learners receive in L2 classrooms. We specifically focused on the relationship between several learner background factors (e.g., L2 contact, native language literacy skills, general academic ability) and measures of L2 speaking by analyzing the data from a large-scale longitudinal project investigating two different methods of teaching English as an L2 (comprehension-based vs. "traditional" L2 program). We found that the learners who had greater contact with English, who showed more interest in reading, whose parents spoke more English, and who attained higher scores in French reading, English vocabulary recognition and general school ability tended to produce more accurate and comprehensible English sentences in an elicited imitation task. However, this relationship was stronger in the comprehension-based than in the traditional program, revealing a possible interaction between type of instruction and learner background variables. Taken together, these findings suggest that certain types of instruction could be more beneficial for learners with certain learning profiles and highlight the importance of investigating interactions between learner background variables and type of instruction in authentic learning contexts.

1. Introduction

The idea that certain types of instruction may be of greater or lesser benefit for learners with different background profiles is not new. A great deal of educational research in the 1970s and 1980s was devoted to the topic of aptitude-treatment interaction (ATI), which explores differences in effectiveness of instruction for learners with various background profiles (for reviews, see Snow et al., 1980; Vatz et al., 2013). At the heart of this research was the idea that an individual's aptitude, defined as "any measurable person characteristic hypothesized to be [required for] successful goal achievement in the treatment(s) studied", may determine this individual's learning outcomes after a particular treatment (Snow, 1991, p. 205). In this sense, the term aptitude is to be understood more broadly than simply intelligence or cognitive capacities underlying learning. Instead, it encompasses a variety of cognitive, personality, experiential, motivational, and social variables, including memory, musical ability, attitudes, motivation, beliefs and many others (Snow, 1991, 1994; Snow and Lohman, 1984).

---

* Corresponding author. Concordia University, Department of Education, 1455 de Maisonneuve Blvd. West, Montréal, Québec, Canada H3G 1M8. Tel.: +1 514 848 2424x2448; fax: +1 514 848 4520.
  E-mail address: pavel.trofimovich@concordia.ca (P. Trofimovich).

A typical finding of ATI research is an interaction between learners' personal profile with respect to a particular variable (e.g., field-dependence) and the type of treatment, such that the effectiveness of treatment differs for learners who differ in their profiles. For example, in one early ATI study, Ridberg et al. (1971) showed that children who differed along the dimension of reflective vs. impulsive behavior responded differently to instructional videos modeling these two behaviors. Compared to reflective children, impulsive children were more strongly influenced by a reflective model, while reflective children were only weakly affected by an impulsive model. In essence, this and other lines of ATI research aimed to determine whether certain learner profiles, understood broadly as aptitudes, were linked to learning with specific materials or in particular contexts and to identify the types of instruction that would match individual learner profiles. However, early attempts to document robust links between learner background profiles and effectiveness of instruction were not altogether successful. Reviewing several decades of ATI research on various aspects of learning, especially mathematics and reading, Snow (1989) reported largely inconclusive findings, which ultimately led to a relative decline in ATI research (see Kalyuga, 2007; for a re-conceptualization of ATI).

In the field of second language (L2) learning, early research on language aptitude was concerned with a narrower cluster of factors within the aptitude complex, namely, cognitive variables believed to underlie the human capacity for language learning (Carroll and Sapon, 1959; Pimsleur, 1966). These variables typically included aspects of short-term memory, phonetic coding or phonological sensitivity (ability to encode and retain auditory sequences), grammatical sensitivity (ability to recognize grammatical functions of words), rote learning ability, imitation or mimicry, and inductive learning (ability to infer rules or patterns from linguistic information). For instance, Pimsleur et al. (1962) examined the relationship between several measures of university-level learners' L2 French proficiency and several learner variables. L2 speaking was shown to be associated with aspects of first language verbal intelligence (which included vocabulary knowledge), motivation, as well as analytical reasoning and rapid articulation of sounds. This early research (see Skehan, 2002), focused on predicting how well an L2 can be acquired by learners with particular cognitive profiles and did not examine whether individual learner factors (such as verbal intelligence or motivation) interact with various types of instruction resulting in different learning outcomes for different learners. In fact, despite initial excitement, empirical work in this area had languished, largely due to the concern that the early research used only cognitive measures of aptitude that were more relevant to traditional, grammar-focused methods of teaching than to language learning in more naturalistic or communicatively oriented contexts (e.g., Krashen, 1982).

The construct of aptitude—in both its narrower sense as cognitive factors underlying language learning and its broader sense as any relevant personal characteristic—is no longer seen as predictive only in traditional teaching methods, and L2 researchers have continued to explore the role of different learner profiles (e.g., individual differences in aptitude, motivation, or attitudes) in L2 learning (see Dörnyei, 2005; Grigorenko et al., 2000; Sparks et al., 2011). For example, Ranta (2002) found that "language analytic ability," a cognitive aspect of aptitude often assumed to be useful for students in grammar-focused instruction, was actually associated with success for young learners in a communicative teaching context, where nearly all activities were meaning-focused. Nevertheless, the question of whether certain types of L2 instruction are particularly effective for learners with specific aptitude profiles has remained unanswered. As discussed in detail below, some reasons for this are that there is still a dearth of research examining this issue, and the few studies on this topic found relatively weak evidence for ATI effects (de Graaff, 1997; Robinson, 1997) or have explored such effects from instructional interventions that are very brief, often spanning only a few contact hours (Erlam, 2005; Robinson, 1997).

Most of the available evidence for ATI effects in L2 learning has come to date from laboratory-based intervention studies. For example, de Graaff (1997) investigated the learning of morphology and syntax in an artificial language by native speakers of Dutch. The learners were exposed to approximately 150 h of self-access computer-based training under either explicit (rules explained) or implicit (rules not explained) conditions. There was little evidence of ATI effects in this study because grammatical sensitivity and rote learning ability had a similar relationship with measures of learning in both the explicit and implicit group. In another study, Robinson (1997) showed that the same two components of aptitude were related to the learning of English grammar rules by L2 learners in three of the four learning conditions investigated (*implicit, rule-search*, and *instructed*, but not *incidental*). Although instructional treatments were brief, involving exposure to 40 sentences, it appears that some aptitude components (grammatical sensitivity, rote learning ability) may be more relevant to certain exposure conditions than others (see also Erlam, 2005). These findings and other research linking various aptitude components to measures of L2 performance (see Robinson, 2012) have led Robinson to propose the Aptitude Complex/Ability Differentiation Model of aptitude (2001, 2012) which posits that different combinations of learner aptitude profiles—defined broadly as clusters of cognitive,

affective, and volitional/attitudinal variables (Snow, 1994)—are associated with learning outcomes in different L2 exposure conditions.

Despite these advances in recent theorizing about aptitude, there is still a dearth of evidence that different kinds of instruction, particularly when such instruction exceeds a few contact hours, are associated with certain learner profiles (see DeKeyser, 2012; Vatz et al., 2013). In a study investigating two adult ESL classes over a six-week period, Spada (1986) found that the amount of contact with the L2 outside class was correlated positively with learners' performance on grammar and writing tests in form-focused classrooms but was associated negatively with these measures when instruction involved less focus on form. In essence, Spada showed that learners with different profiles of L2 contact outside the classroom tended to benefit from different kinds of instruction inside the classroom. Similar interactions were reported by Flege and Liu (2001), with length of residence used as a measure of L2 experience, by Tang (1971), who found that the effectiveness of different types of instruction was associated with learner attitudes toward native (L1) speakers, and Niwa (2000) showing that aspects of aptitude and intelligence are more strongly associated with learner performance on complex than simple learning tasks.

Clearly, documenting possible ATI effects (and extending such effects beyond purely cognitive components of aptitude, such as grammatical sensitivity or rote learning ability) requires longitudinal investigations that not only compare learners following distinct instructional programs but that also include a range of learner background measures. To address this goal, we investigated possible ATI effects in a large dataset from a two-year longitudinal comparison of L2 learners in two contexts (comprehension-based vs. "traditional" L2 programs). The two programs differed dramatically in the input opportunities they provided to students, with the comprehension-based program engaging students only in listening and reading activities and the traditional program featuring practice in all four language skills. For these learners, we examined several background variables (e.g., L2 contact, L1 literacy skills, general academic ability) in relation to two measures of learners' L2 speaking ability (accuracy, comprehensibility) in each instructional context.[1] An emphasis on L2 speaking (i.e., a skill that was overtly practiced in the traditional but not in the comprehension-based program) was intentional, in that it allowed us to determine whether a different set of learner profile variables was associated with the performance in the skill that was either supported or unsupported through instruction. Thus, the two programs, with their different emphasis on speaking practice, provided an interesting longitudinal context for examining ATI effects, namely, investigating whether the relationship between learner background variables and measures of L2 speaking depends on the instructional context.

Therefore, the goal of the current study was to investigate a possible link between L2 learners' background variables and the type of instruction learners receive in L2 classrooms. In particular, our analysis was guided by the following question: Does the relationship between learner background variables (e.g., L2 contact, L1 literacy skills, general academic ability) and measures of their L2 speaking ability depend on the nature of instruction received by the learners? We hypothesized that, consistent with ATI effects, the two instructional contexts might draw on different learner variables, revealing different sets of relationships between learner background variables and learning outcome measures. Our overall objective in carrying out these analyses was to highlight the importance of investigating ATI effects in authentic learning contexts and to consider the relevance of such research to L2 pedagogy.

2. Method

Our data come from a large-scale longitudinal study of young francophone learners of English in New Brunswick, Canada. The participants were in either a traditional (henceforth "regular") program or a new comprehension-based (henceforth "experimental") program (Lightbown, 1992). The current dataset included several longitudinal measures of speaking ability from Trofimovich et al. (2009) for a sample of 73 grade 3 ESL learners, of whom 48 were enrolled in the experimental and 25 in the regular program (M_{age} = 8.4 years). The learners, drawn randomly from 20 intact classes following each type of program, had been brought up in francophone families, had resided in francophone areas, and had received no English instruction prior to grade 3.

---

[1] Although the measures reported here as part of learner background profiles were restricted to already existing data and may have thus represented a narrow sample of all possible aptitude variables, the current dataset nevertheless spans a broad range of aptitude characteristics (Snow, 1994) and arguably features a set of learner profile factors relevant to both the target context and a variety of other similar environments.

The regular ESL program was a modified aural-oral program taught by L2 speakers of English with little or no specialized ESL training. In the regular program, learners were typically engaged in question-and-answer activities, dialogues, songs, and a minimal amount of reading and writing practice. The experimental program was created in 1985 to alleviate the shortage of qualified ESL teachers. Activities in this program were entirely comprehension-based and took place in classrooms equipped with a cassette player and a headset for each learner and stocked with a large collection of printed materials with accompanying cassette tapes. The materials included picture/story books designed for young English-speaking children, some graded readers designed for L2 learners, picture dictionaries, age-appropriate ESL textbooks, and simple science books and biographies. During each class, learners worked on their own, reading and listening to the materials they selected for that day. Thus, unlike the regular program, the experimental program included no lessons, tests, interaction with other learners, or feedback from teachers. Students worked individually and independently for the 30-min period of their daily ESL class (see Lightbown, 1992; Lightbown et al., 2002; for further description of the classroom experiences of learners in each program).

The learners in both programs were pretested at the beginning of grade 3, and the two groups did not differ in terms of several background variables (all rated by the children's parents on a written questionnaire): daily contact with English (−7.9 vs. −7.8 on a scale between −14 = *no contact with English* and +14 = *lots of contact with English*), interest in reading in English (−3.3 vs. −3.1 on a scale between −7 = *no interest* and +7 = *great interest*), and the extent to which parents themselves were proficient in English (−.4 vs. .3 on a scale between −6 = *limited proficiency* and +6 = *nativelike*). Similarly, the learners in the experimental and regular groups did not differ in their performance on several tests: the English aural vocabulary recognition (AVR) test (23.8 vs. 24.3, range = 8−64), the French reading test (7.9 vs. 8.0, range = 5.0−10.5), and the French version of the Otis−Lennon school ability test (53.8 vs. 52.2, range = 23−69). The AVR test (which required children to match pictures to spoken words heard on tape) was administered at the beginning of grade 3. The French reading test, a province-wide French reading exam, was administered at the end of grade 1.[2] The Otis−Lennon test, a test of general academic ability, was administered at the end of grade 3.

Learners' English speaking ability was assessed twice (at the end of grade 3 and then again at the end of grade 4) by means of an elicited imitation activity in which learners repeated six simple sentences (played back on a cassette tape one at a time). The sentences, recorded by a female English speaker, were spoken slowly and clearly and sounded like a short story read aloud to young children. Upon hearing each sentence (e.g., *His dog's name is Tom*), learners attempted to repeat it to the best of their ability (e.g., *His dog name Tom*). Students were tested individually and all their responses were audio recorded for later analysis.

Two of the measures derived from the elicited imitation task (see Trofimovich et al., 2009; for details of how student response data were prepared for analysis) were used in the current analyses, although similar findings were obtained for other measures used in previous reports. These two measures—accuracy and comprehensibility of L2 elicited imitation—tapped into two distinct aspects of L2 speaking ability. Accuracy of elicited imitation was designed to be an implicit measure of learners' ability to produce L2 speech, that is, a measure that depended little on learners' metalinguistic knowledge of language (Ellis, 2005). This measure was calculated as the number of words from each original sentence prompt that were correctly reproduced (i.e., with all morphemes intact) over the total number of words in each sentence prompt. This measure was scored by phonetically trained judges (interrater reliability = .98). In turn, comprehensibility was selected as an index of learners' communicative effectiveness, since comprehensibility is a measure of listener effort in understanding L2 speech (Munro and Derwing, 1999). This measure was defined as the average of ratings by 20 native English speakers who listened to each learner's six sentences (one at a time) and rated how easily they could understand each sentence on a 9-point scale (1 = *hard to understand*, 9 = *easy to understand*; interrater reliability = .90−.96). To summarize, our main interest was to investigate whether the relationship between measures of L2 elicited imitation accuracy and comprehensibility on the one hand, and individual learner profiles (defined as learner contact with English, reading interest, parental bilingualism, English vocabulary recognition performance, and scores in grade 1 French reading and general learning ability) on the other, interacted with the instructional context that placed different emphasis on L2 speaking practice.

---

[2] The data for the French reading test were made available to the researchers by the Ministry of Education of New Brunswick for all learners participating in this project. Because this test was a province-wide examination administered in grade 1, it was impossible to obtain equivalent literacy data at the beginning of the study (i.e., for grade 3 learners).

3. Results

Briefly, the results of between-group comparisons (reported in full in Trofimovich et al., 2009) showed that the regular and the experimental groups both improved over time. For *accuracy*, there was no significant difference between the groups. For *comprehensibility*, however, the regular group received slightly higher ratings than the experimental group. Based on these findings, Trofimovich et al. (2009) argued that at least two years of comprehension-based experience, in the absence of any speaking practice inside the classroom, was not detrimental to the development of L2 pronunciation, but noted some possible limits of comprehension-only experience in promoting fluent and clearly articulated L2 speech.

The remaining analyses focus on the main goal of the current report, namely, investigating possible interactions between instructional and learner background variables. The research question guiding this study was whether the relationship between learner background variables (e.g., L2 contact, L1 literacy skills, general academic ability) and measures of their L2 speaking ability depends on the nature of instruction received by the learners. To address this question, Pearson correlation analyses were carried out separately for the learners in the two programs. In each analysis, associations were computed between the background variables on the one hand and accuracy scores and comprehensibility ratings at the end of grades 3 and 4 on the other. Tables 1 and 2 show the results of these analyses.

For the experimental group (Table 1), most background measures (contact with English, reading interest, parental bilingualism), initial proficiency (AVR score), and general learning ability (grade 1 French reading, Otis–Lennon) correlated positively, revealing moderate-strength associations, with the outcome measures at the end of grade 3 and even more markedly at the end of grade 4. In grade 3, these associations mostly involved learners' comprehensibility scores. In grade 4, similar associations obtained for both the accuracy and comprehensibility measures, suggesting that learner background variables were less likely to be linked to speaking accuracy in the first than in the second year of the program. In contrast, although several moderate-strength associations were found for the regular group (Table 2), they were few in number and were found only between the background variables and the comprehensibility measure.

This finding is noteworthy given that there were no statistically significant differences between the experimental and regular groups for the three background measures that were re-administered at the end of grade 4, namely, contact with English, reading interest, and AVR scores based on another version of the test, $ts < 1.33$, $ps > .19$. In other words, the learners in both groups were highly comparable at the end of the study in terms of these profiles, yet how these profiles related to their L2 speaking accuracy and comprehensibility depended on the instruction the students had received.

We conducted two additional analyses to rule out possible alternative explanations. First, we examined whether the patterns of correlations were affected by the nature and the distribution of the data. Goodwin and Leech (2006) identify several factors influencing the strength of correlation coefficients; among these are data variability, presence of outliers, and sample size. A comparison of the data for the six background measures (contact with English, reading interest, parental bilingualism, AVR score, grade 1 French reading, Otis–Lennon) showed that there were no marked differences in the distribution of scores between the regular and the experimental groups. Levene's tests for equality of variances revealed no significant differences between the two groups for any of the background measures, $Fs < 3.52$, $ps > .07$, except for parental bilingualism, with the experimental group characterized by a larger variance than the regular group, $F = 11.72$, $p > .001$. Similarly, one-sample Kolmogorov–Smirnov tests showed that the distributions for all scores were normal in both groups, $Zs < 1.43$, $ps > .07$, apart from reading interest scores in the

Table 1
Pearson correlations between learner background variables and grade 3 and 4 performance in the experimental program.

Variable	Accuracy		Comprehensibility	
	Grade 3	Grade 4	Grade 3	Grade 4
Contact with English	.22	.32*	.41**	.32*
Reading interest	.13	.26	.34*	.35*
Parental bilingualism	.32*	.38**	.34*	.48***
Gr. 1 French reading	.41**	.42**	.41**	.57***
AVR score	.22	.30*	.48***	.46***
Otis–Lennon	.09	.31*	.18	.36*

Note. *$p < .05$, **$p < .01$, ***$p < .001$ (two-tailed).

Table 2
Pearson correlations between learner background variables and grade 3 and 4 performance in the regular program.

Variable	Accuracy		Comprehensibility	
	Grade 3	Grade 4	Grade 3	Grade 4
Contact with English	.16	−.06	.45*	.24
Reading interest	.04	.11	.18	.25
Parental bilingualism	.03	.24	.33	.49*
Gr. 1 French reading	.29	.18	.04	.14
AVR score	.18	.26	.52**	.56**
Otis−Lennon	.22	.15	−.06	.01

Note. *$p < .05$, **$p < .01$ (two-tailed).

experimental group, which featured a positive skew (i.e., indicative of low interest in reading). In addition, none of the distributions contained any obvious outlier scores which could have influenced the size of correlation coefficients. Finally, the obtained pattern of correlations was unlikely to have been caused by differences in sample size between the two groups, as sample size usually affects the significance associated with the obtained correlations, not their strength (Goodwin and Goodwin, 1999; Goodwin and Leech, 2006). To summarize, we found no obvious statistical explanations for the obtained pattern of associations.

In the second follow-up analysis, we examined whether the obtained pattern of findings was caused by Matthew effects. Briefly, Matthew effects are defined as "the amplification of any initial advantage (e.g., economic resources, health status, cognitive ability) that leads to cumulative differences that widen preexisting gaps" (Ceci and Papierno, 2005, p. 149). For example, research on reading has often shown that readers with larger vocabularies tend to read extensively, increasing their knowledge base, while readers for whom reading comprehension is problematic because their vocabulary knowledge is low are likely to avoid reading and as a result will keep falling behind. In essence, the impact of reading practice appears to interact with learners' initial level of vocabulary knowledge (Stanovich, 1986).

Could the relationship between type of instruction and learner background variables in the experimental group reflect Matthew effects, such that the experimental program benefits only learners with stronger background profiles and has little or even negative influence on weaker learners? To answer this question, we examined individual learner performance trajectories over time. Fig. 1 illustrates such trajectories plotted for accuracy. The performance trajectories for the experimental group resemble a fan-shaped pattern typical of Matthew effects: learner performance is "tighter" in grade 3 but is more "spread out" and variable in grade 4. However, a closer inspection of the data revealed that the students with higher scores in grade 3 did not necessarily improve and that those with poorer scores did not necessarily fall behind. In fact, we found an approximately equal number of initially low- and high-performing students who eventually improved (31% vs. 21%) or regressed in their performance (21% vs. 27%).

The comprehensibility data for the experimental group (Fig. 2) were even less consistent with Matthew effects. There, improvement across time was unrelated to initial performance levels: 42% of initially low- and 44% of initially high-performing students improved by the end of grade 4. These findings suggest that the experimental program benefitted learners regardless of the level of comprehensibility they had achieved in grade 3.
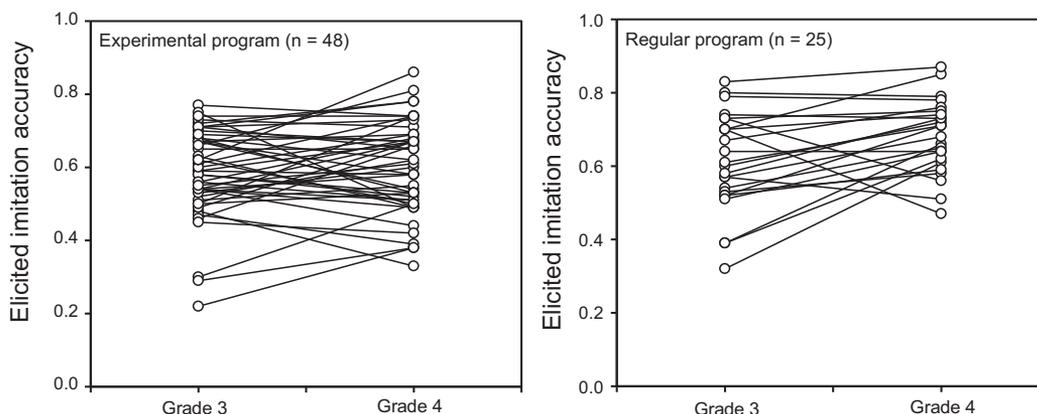

Fig. 1. Individual longitudinal performance trajectories (accuracy) for learners in the experimental and regular programs.
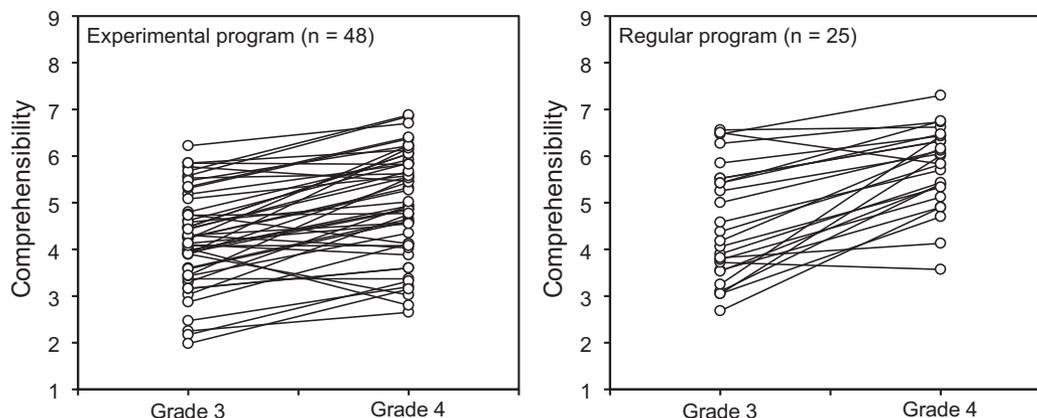
Fig. 2. Individual longitudinal performance trajectories (comprehensibility) for learners in the experimental and regular programs.

4. Discussion

To summarize, we found that the students who had greater contact with English, who showed more interest in reading, whose parents spoke more English, and who attained higher scores in French reading, English vocabulary recognition and general school ability tended to produce more accurate and comprehensible English sentences in an elicited imitation task. However, this relationship was stronger in the experimental than in the regular program, which showed a possible interaction between type of instruction and measures of learners' L2 contact, L1 literacy, and learning ability. These associations for the experimental group held across time and a range of achievement measures. For example, learners' performance on the French reading test (administered in grade 1) and on the Otis−Lennon school ability test (administered in grade 3) had similar positive correlations with grade 4 measure of accuracy and grade 3 and 4 measures of comprehensibility.

We also showed that the stronger correlations between learner background variables and achievement measures in the experimental group were not due to Matthew effects, which would imply that the experimental program favored only "better" students and disadvantaged weaker ones. In fact, in the experimental group, there appeared to be no disadvantage for students with less ability or prior knowledge, as they seemed to learn as much as they would have through regular instruction (Fig. 1). However, as our correlation analyses indicated (Table 1), there were genuine advantages for students with stronger learning profiles who were prepared to learn more and to take advantage of what was on offer, namely, large amounts of high-quality input in a learning environment that encouraged autonomous learning without any pressure to perform.

One possibility to interpret the interactions between learner background variables and type of instruction in this study, and possibly ATI effects found in prior research, is to suggest that they are typical of learning environments that involve a certain degree of transferability of skills from a language classroom to the "real world" (Lightbown, 2007). From the standpoint of cognitive psychology such learning environments often feature cognitive skills typical of real-life situations of language use (in this case, listening to input spoken by many different speakers). These learning contexts also often challenge learners to work harder to accomplish a learning goal, with an underlying assumption that more effortful processing at the time of learning leads to longer-lasting, stronger memories needed for efficient use of the learned information (Bjork, 1994).

The experimental program in this study is perhaps one such challenging learning context, featuring rich real-life language input directly transferable to real-world language use. Therefore, all learners in the experimental group benefited from the comprehension-based program; in addition, the program likely provided further advantages for learners with initially stronger learning profiles (with greater prior knowledge, extent of L2 exposure, L1 literacy, and learning ability). In contrast, the regular program likely offered limited content and featured a small repertoire of language patterns and skills that did not easily transfer outside the classroom. Consequently, such a learning environment may have failed to fully exploit learners' strengths in terms of individual differences. This likely led to a leveling off process in which learners with the potential to learn more did not.

Although it may be tempting to interpret interactions between learner background variables and type of instruction in terms of transferability of learning skill (see Sternberg et al., 1999; for preliminary evidence), this view clearly

warrants further research. At least one goal of this research would be to isolate the learner profiles that are most relevant to the transfer of learning outside the classroom in a given learning context. Another avenue of future research is to investigate possible bi-directional, reciprocal links between learner-background variables and instructional effects. On the one hand, as we have argued above, certain types of instruction likely draw on certain learner background variables, allowing initially stronger learners to gain additional benefits from instruction. Alternatively, particular types of instruction could also positively impact certain individual learner variables, thus allowing learners, for example, to increase their L2 exposure, improve their L1 literacy skills, and possibly even enhance their cognitive processing capacity. Being able to harness such positive, self-reinforcing loops for pedagogical purposes might indeed open possibilities for creating new and efficient ways of L2 instruction.

To conclude, in a longitudinal investigation of two dramatically different instructional programs, we have shown strong interactions between learner background profiles and type of instruction, with learner background variables associated with achievement measures in the comprehension-based program. These findings highlight the importance of investigating ATI effects and similar interactions involving other individual learner differences in authentic learning contexts, with a view to developing instructional treatments and designing entire teaching programs that would allow all learners, including those with stronger and weaker background profiles, to benefit from instruction. Most likely, such instructional treatments and programs would involve a certain degree of learning challenge and capitalize on those skills that permit learners to transfer what they have learned to contexts of real-life language use.

References

Bjork, R.A., 1994. Memory and metamemory considerations in the training of human beings. In: Metcalfe, J., Shimamura, A. (Eds.), Meta-cognition: Knowing about Knowing. MIT Press, Cambridge, MA, pp. 185–205.

Carroll, J.B., Sapon, S., 1959. The Modern Language Aptitude Test. Psychological Corporation, San Antonio, TX.

Ceci, S.J., Papierno, P.B., 2005. The rhetoric and reality of gap closing: when the "have-nots" gain but the "haves" gain even more. Am. Psychol. 60, 149–160.

de Graaff, R., 1997. The Experanto experiment: effects of explicit instruction on second language acquisition. Stud. Second Lang. Acquisition 19, 249–276.

DeKeyser, R., 2012. Interactions between individual differences, treatments, and structures in SLA. Lang. Learn. 62, 189–200.

Dörnyei, Z., 2005. The Psychology of the Language Learner: Individual Differences in Second Language Acquisition. Lawrence Erlbaum, Mahwah, NJ.

Ellis, R., 2005. Measuring implicit and explicit knowledge of a second language. A psychometric study. Stud. Second Lang. Acquisition 27, 141–172.

Erlam, R., 2005. Language aptitude and its relationship to instructional effectiveness in second language acquisition. Lang. Teach. Res. 9, 147–171.

Flege, J.E., Liu, S., 2001. The effect of experience on adults' acquisition of a second language. Stud. Second Lang. Acquisition 23, 527–552.

Goodwin, L.D., Goodwin, W.L., 1999. Measurement myths and misconceptions. Sch. Psychol. Q. 14, 408–427.

Goodwin, L.D., Leech, N.L., 2006. Understanding correlation: factors that affect the size of r. J. Exp. Education 74, 251–266.

Grigorenko, E., Sternberg, R.J., Ehrman, M.E., 2000. A theory based approach to the measurement of foreign language learning ability: the Canal-F theory and test. Mod. Lang. J. 84, 390–405.

Kalyuga, S., 2007. Expertise reversal effect and its implications for learner-tailored instruction. Educ. Psychol. Rev. 19, 509–539.

Krashen, S.D., 1982. Principles and Practice in Second Language Acquisition. Pergamon, Oxford.

Lightbown, P.M., 1992. Can they do it themselves? A comprehension-based ESL course for young children. In: Courchêne, R., St John, J., Thérien, C., Glidden, J.I. (Eds.), Comprehension-based Second Language Teaching. University of Ottawa Press, Ottawa, pp. 353–370.

Lightbown, P., 2007. Transfer appropriate processing as a model for classroom second language acquisition. In: Han, Z. (Ed.), Understanding Second Language Process. Multilingual Matters, Clevedon, pp. 27–44.

Lightbown, P.M., Halter, R.H., White, J., Horst, M., 2002. Comprehension-based learning: the limits of "do it yourself." Can. Mod. Lang. Rev. 58, 427–464.

Munro, M.J., Derwing, T.M., 1999. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. Lang. Learn. 49, 285–310.

Niwa, Y., 2000. Reasoning Demands of L2 Tasks and L2 Narrative Production: Effects of Individual Differences in Working Memory, Intelligence and Aptitude. Unpublished MA thesis. Department of English, Aoyama Gakuin University, Japan.

Pimsleur, P., 1966. The Pimsleur Language Aptitude Battery. Harcourt, Brace, Jovanovich, New York.

Pimsleur, P., Stockwell, R.P., Comrey, A.L., 1962. Foreign language learning ability. J. Educ. Psychol. 53, 15–26.

Ranta, L., 2002. The role of learners' language analytic ability in the communicative classroom. In: Robinson, P. (Ed.), Individual Differences and Instructed Language Learning. John Benjamins, Amsterdam, pp. 159–180.

Ridberg, E.H., Parke, R.D., Hetherington, E.M., 1971. Modification of impulsive and reflective cognitive styles through observation of film-mediated models. Dev. Psychol. 5, 369–377.

Robinson, P., 1997. Individual differences and the fundamental similarity of implicit and explicit adult second language learning. Lang. Learn. 47, 45–99.

Robinson, P., 2001. Individual differences, cognitive abilities, aptitude complexes and learning conditions in second language acquisition. Second Lang. Res. 17, 368–392.

Robinson, P., 2012. Individual differences, aptitude complexes, SLA processes and aptitude test development. In: Pawlak, M. (Ed.), New Perspectives on Individual Differences in Language Learning and Teaching. Springer, Berlin, pp. 57–75.

Skehan, P., 2002. Theorising and updating aptitude. In: Robinson, P. (Ed.), Individual Differences and Instructed Language Learning. John Benjamins, Amsterdam, pp. 69–95.

Snow, R.E., 1991. Aptitude-treatment interaction as a framework for research on individual differences in psychotherapy. J. Consult. Clin. Psychol. 59, 205–216.

Snow, R.E., 1994. Abilities in academic tasks. In: Sternberg, R., Wagner, R. (Eds.), Mind in Context: Interactionist Perspective on Human Intelligence. Cambridge University Press, New York, pp. 3–37.

Snow, R.E., Lohman, D.F., 1984. Toward a theory of cognitive aptitude for learning from instruction. J. Educ. Psychol. 76, 347–376.

Snow, R., 1989. Aptitude-treatment interaction as a framework for research on individual differences in learning. In: Ackerman, P.L., Sternberg, R.J., Glaser, R. (Eds.), Learning and Individual Differences: Advances in Theory and Research. W.H. Freeman, New York, pp. 13–59.

Snow, R., Federico, P., Montague, W., 1980. Aptitude, Learning, and Instruction. Lawrence Erlbaum, Hillsdale, NJ.

Spada, N., 1986. The interaction between type of contact and type of instruction: some effects on the L2 proficiency of adult learners. Stud. Second Lang. Acquisition 8, 181–200.

Sparks, R.L., Humbach, N., Patton, J., Ganschow, L., 2011. Subcomponents of second-language aptitude and second-language proficiency. Mod. Lang. J. 95, 253–273.

Stanovich, K.E., 1986. Matthew effects in reading: some consequences of individual differences in the acquisition of literacy. Reading Res. Q. 21, 360–407.

Sternberg, R.J., Grigorenko, E.L., Ferrari, M., Clinkenbeard, P., 1999. A triarchic analysis of an aptitude-treatment interaction. Eur. J. Psychol. Assess. 15, 3–13.

Tang, B.S.T., 1971. A Psycholinguistic Study of the Relationship between Children's Ethnic-linguistic Attitudes and the Effectiveness of Methods Used in Second-language Reading Instruction. Unpublished doctoral dissertation. Stanford University, Stanford, CA.

Trofimovich, P., Lightbown, P.M., Halter, R., Song, H., 2009. Comprehension-based practice: the development of L2 pronunciation in a listening and reading program. Stud. Second Lang. Acquisition 31, 609–639.

Vatz, K., Tare, M., Jackson, S.R., Doughty, C.J., 2013. Aptitude-treatment interaction studies in second language acquisition: findings and methodology. In: Granena, G., Long, M. (Eds.), Sensitive Periods, Language Aptitude, and Ultimate L2 Attainment. John Benjamins, Amsterdam, pp. 273–292.