Studies in Second Language Acquisition, 2011, **33**, 589–624.
doi:10.1017/S0272263111000325

LEARNED ATTENTION IN ADULT LANGUAGE ACQUISITION

A Replication and Generalization Study and Meta-Analysis

Nick C. Ellis
University of Michigan

Nuria Sagarra
Pennsylvania State University

This study investigates associative learning explanations of the limited attainment of adult compared to child language acquisition in terms of learned attention to cues. It replicates and extends Ellis and Sagarra (2010) in demonstrating short- and long-term learned attention in the acquisition of temporal reference in Latin. In Experiment 1, salient adverbs were better learned than less salient verb inflections, early experience of adverbial cues blocked the acquisition of verbal morphology, and, contrariwise—but to a lesser degree—early experience of tense reduced later learning of adverbs. Experiment 2 demonstrated long-term transfer: Native speakers of Chinese (no tense morphology) were less able than native speakers of Spanish or Russian (rich morphology) to acquire inflectional cues from the same language experience where adverbial and verbal cues were equally available. Learned attention to tense morphology in Latin was continuous rather than discrete, ordered with regard to first language: Chinese

This research was supported by National Science Foundation grant BCS-0717557. We thank Katherine Martin for help designing, piloting, scoring, and administering these experiments. The first author is grateful for the support of Freiburg Institute for Advanced Studies (FRIAS) where he was based as an External Senior Fellow while revising this article. Nuria Sagarra is now at Carnegie Mellon University.

Address correspondence to: Nick C. Ellis, English Language Institute, University of Michigan, 500 E. Washington Street, Ann Arbor, MI 48104; e-mail: ncellis@umich.edu.

© Cambridge University Press 2011

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

< English < Russian < Spanish. A meta-analysis of the combined results of Ellis and Sagarra and the current study separates out positive and negative learned attention effects: The average effect size for entrenchment was large (+1.23), whereas that for blocking was moderate (–0.52).

Naturalistic foreign language acquisition (FLA) tends to stabilize at levels short of nativelike ability. At its most extreme, this can present itself as a basic variety of interlanguage that, although sufficient for everyday communicative purposes, predominantly comprises just nouns, verbs, and adverbs, with closed-class items—in particular, grammatical morphemes and prepositions—failing to be put to full nativelike use (Bardovi-Harlig, 1992; Clahsen & Felser, 2006; Klein & Rieck, 1982, cited in Klein, 1986; Mangubhai, 1991; Schmidt, 1984; VanPatten, 1996, 2006). Foreign language (FL) learners initially make temporal references mostly by use of adverbials, prepositional phrases, serialization, and calendric reference, with the grammatical expression of tense and aspect emerging only slowly thereafter, if at all (Bardovi-Harlig, 1992, 2000; Lee, 2002; Meisel, 1987; Noyau, Klein, & Dietrich, 1995; Sagarra, 2001).

One factor determining cue selection is salience: Prepositional phrases, temporal adverbs, and other lexical cues to time are quite pronounced in the speech stream. Verbal inflections are not (consider ***yesterday*** *I walk****ed***). The low salience and low reliability of grammatical cues tends to make them less learnable (Ellis, 2006c; Goldschneider & DeKeyser, 2001) and could underlie late learners' difficulty in processing and producing FL verbal morphology (Jiang, 2004; Zobl & Liceras, 1994). However, salience and reliability affect first language (L1) acquisition and FLA alike—there has to be something else that accounts for the limitations in FLA.

Associative learning theory documents a range of effects of transfer and inhibition that shift learners' attention to input as a result of prior experience. Kamin (1969) and Kruschke (2006) described the phenomenon of blocking. Learning that a particular stimulus is associated with a particular outcome makes it harder to learn that another cue, subsequently paired with that same outcome, is also a good predictor of it. For example, if an animal learns that a conditioned stimulus is a reliable predictor of an unconditioned stimulus (e.g., that a light reliably predicts the onset of some painful stimulus such as a shock), then it will not become conditioned to or learn that any other conditioned stimulus predicts that unconditioned stimulus (e.g., that a bell predicts the onset of the shock the same way the light did). The prior association essentially blocks further associations. Blocking is an effect of learned attention (Kamin; Kruschke & Blair, 2000; Mackintosh, 1975). It is a highly

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

robust and widespread phenomenon, occurring across animal and human learning (Rescorla & Wagner, 1972; Shanks, 1995; Wills, 2005).

Ellis (2006c) reviews the phenomenon as it might apply in SLA. There are many situations in natural language where cues are redundant (Schmidt, 2001; Terrell, 1991; VanPatten, 1996) and thus, as a consequence of blocking, might be less readily learned. If a learners' L1 experience has led them to look elsewhere for cues to interpretation, they might use these cues where available in the FL—and if they do, the principles of associative learning predict that this will be to the detriment of learning other cues that might also be relevant. For example, L1-derived knowledge that there are reliable lexical cues to temporal reference (words like *gestern, hier, ayer, yesterday*) might block the acquisition of verb tense morphology from analysis of utterances such as *yesterday I walked*.

Various theories of SLA incorporate related notions of transfer and learned attention. The competition model (MacWhinney, Bates & Kliegl, 1984; MacWhinney & Bates, 1989; MacWhinney, 2001) was explicitly formulated to deal with competition between multiple linguistic cues to interpretation. Input processing theory (VanPatten, 1996) includes the lexical preference principle: "Learners will process lexical items for meaning before grammatical forms when both encode the same semantic information" (VanPatten, 2006, p. 118). This principle encapsulates the mounting evidence that FL learners prefer lexical to grammatical cues (for oral tasks, see Lee, Cadierno, Glass, & VanPatten, 1997; Musumeci, 1989; for written tasks, see Lee, 1999). The preference for nonredundancy principle is also included in this theory: "Learners are more likely to process nonredundant meaningful grammatical markers before they process redundant meaningful markers" (VanPatten, 2006, p. 119). The associative-cognitive CREED (whereby SLA is construction-based, rational, exemplar-driven, emergent, and dialectic; Ellis, 2006a, 2006b, 2006c, 2008b) describes the limited end state typical of FLA directly in terms of learned attention, salience, overshadowing, and blocking.

Ellis and Sagarra (2010) explored these phenomena of learned attention in two experiments. The first demonstrated short-term instructional sequence effects in adults learning temporal reference in Latin using the standard blocking experimental paradigm (Kruschke, 2006) but with linguistic content. Native English speakers learned a small number of Latin expressions and their English translations. There were three groups: adverb pretraining, control, and verb pretraining. In Phase 1, adverb pretraining participants learned two adverbs and their temporal reference: *hodie* "today" (present) and *heri* "yesterday" (past); verb pretraining participants learned two verbs and their temporal reference: *cogito* "I think" (present) and *cogitavi* "I thought" (past); the control group had no such pretraining. In Phase 2, all participants were shown sentences that appropriately combined an adverb and a verb

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

(e.g., *heri cogitavi* "yesterday I thought," *hodie cogito* "today I think," *cras cogitabo* "tomorrow I will think") and learned whether these sentences referred to the present, the past, or the future. In Phase 3 (the reception test), all combinations of adverb (*hodie* "today," *heri* "yesterday," *cras* "tomorrow") and verb tense marking (*cogit**o*** "I think," *cogit**avi*** "I thought," *cogit**abo*** "I will think") were combined and the participants were asked to judge, without feedback, whether each sentence referred to the past, present, or future on a scale of 1 (*past*) to 5 (*future*). The logic of the design was as follows. In Phase 2, every utterance contained two temporal references—an adverb and a verbal inflection. If participants paid equal attention to these two cues, then in Phase 3, their judgments should be equally affected by both the adverb and the verb. If, however, participants paid more attention to adverbial (or verbal) cues, then their judgments would be swayed toward these cues in Phase 3.

The results showed that the three groups reacted to the cues in very different ways. In two-word sentences, there was temporal information cued by both an adverb and an inflection; when these cues deviated, the verb pretraining group followed the verbal cue, the adverb pretraining group followed the adverbial cue, and the control group lay in between. Multiple regression analyses—one for each group, where the dependent variable was group mean temporal interpretation for each of the Phase 3 strings and the independent variables were the information conveyed by the adverbial and verbal inflection cues—estimated the differential cue use by each of the three groups, in standardized β coefficients, as shown in Table 1.

An additional design element allowed the inference that these results illustrate attentional biases to particular dimensions of cue (adverb vs. verbal inflection) rather than to particular words. In Phase 1, participants in the adverb and verb pretraining groups learned particular constructions relating to the present and the past. There was no reference to future at this stage. Thus, whereas subsequent responses relating to past and present judgments could reflect specific prior-learned associations, responses relating to future judgments could not. The Phase 3 results showed that the adverb and verb pretraining groups were as unalike and dissociated in their performance on *cras* "tomorrow" and *cogitabo* "I will think" items referring to the future as they were on the other past and present reference ones.

Table 1. Differential cue use across groups in Ellis and Sagarra (2010)

Group	Time	Adjusted R^2
Adverb pretraining	0.97 Adverb + 0.23 Verb	0.98
Verb pretraining	0.12 Adverb + 0.97 Verb	0.96
Control	0.60 Adverb + 0.72 Verb	0.85

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

Ellis and Sagarra's (2010) Experiment 2 illustrated long-term language transfer effects whereby the nature of learners' L1 (± verb tense morphology) biased the acquisition of morphological versus lexical cues to temporal reference in the same subset of Latin. Native speakers of Chinese (whose L1 does not exhibit verb tense morphology) learning under the control conditions of Experiment 1 produced Phase 3 performance reflecting cue use of 0.91 Adverb + 0.29 Verb, adjusted R^2 = 0.90. Comparing these results with those from Experiment 1, it can be seen that they lie closer to those of the original adverb group rather than the original control group. These findings suggest a long-term attention to language, a processing bias affecting subsequent cue learning that comes from a lifetime of prior L1 usage.

RESEARCH QUESTIONS

The current research aims to extend these findings, replicating the basic design but with a more typical, complete, and complicated verbal paradigm. In Ellis and Sagarra (2010), there was just one verb form (first-person singular) for each tense, a selection that artificially balanced the complexity and salience of the adverbial and inflectional cues. Experiment 1 of the current study therefore extends the sample of verbal inflections to more adequately reflect natural language. This manipulation allows determination of the replicability of these earlier demonstrations of learned attention—whether early experience of adverbial cues blocked the acquisition of verbal tense morphology, and, contrariwise, early experience of tense blocked later learning of adverbs. It also allows examination of whether, under control conditions devoid of pretraining, verbal inflections in themselves are less readily acquired than adverbial cues, either as a result of their lower salience, greater complexity, or overshadowing by adverbial cues. Experiment 2 extends the investigation of long-term learned attention effects to L1s with wider variation in use of inflectional morphology. It investigates (a) whether L1 speakers of Chinese (no verb tense morphology) are less able than L1 speakers of Spanish or Russian (rich morphology) to acquire inflectional cues from the same language experience where adverbial and verbal cues are equally available, and (b) whether learned attention to inflectional tense morphology in Latin is continuous rather than discrete, ordered with regard to increasing L1 degree and means of use of morphological cuing of tense: Chinese < English < Russian < Spanish.

EXPERIMENT 1

Experiment 1 involves the learning of a small number of Latin expressions and their English translations. It investigates the effects of successive

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

learning of different types of cues for temporal reference, adverbs (*hodie* "today," *heri* "yesterday," *cras* "tomorrow") and verbal inflections in three tenses and three persons (*cogito*/*cogitas*/*cogitat* "I/you/he think(s)," *cogitavi*/*cogitavisti*/*cogitavit* "I/you/he thought," *cogitabo*/*cogitabis*/*cogitabit* "I/you/he will think"). It determines if the acquisition of one set of cues is impaired if another is already known as a reliable indicator of event time.

Participants

Fifty participants took part in this experiment; 45 were students from a major university in the United States and 5 were other adults from the local community. All were monolingual English speakers. They were volunteers and were paid $10 for their participation in the experiment. None had learned Latin previously. They were randomly allocated to one of three groups: the adverb pretraining group of 5 males and 11 females, age range 18–24 years (mean: 20.0; median: 19.5); the verb pretraining group of 5 males and 12 females, age range 19–31 years (mean: 21.7; median: 21.0); or the no pretraining control group of 5 males and 12 females, age range 18–22 years (mean: 19.6; median: 19.0).

Procedure

The experiment was programmed in E-Prime (Schneider, Eschman, & Zuccolotto, 2002). It took less than 1 hr and comprised four phases: pretraining (Phase 1), sentence decoding (Phase 2), reception testing (Phase 3), and production testing (Phase 4). The procedure of the first three phases is schematized in Figure 1.

Phase 1. Participants in the adverb pretraining group had 36 randomized trials in which they saw either the adverb *hodie* "today" or the adverb *heri* "yesterday." The participants had to choose whether the correct translation was *today* or *yesterday* by clicking on the appropriate alternative with the mouse. These alternatives appeared in counterbalanced positions on the screen. A correct choice returned the feedback "correct," whereas an incorrect one returned, for example, "wrong—the meaning of *hodie* is *today*."

Participants in the verb pretraining group had 36 trials of training on verb inflections rather than adverbs. On each trial they saw one of the past (*cogitavi* "I thought," *cogitavisti* "you thought," *cogitavit* "he thought") or present (*cogito* "I think," *cogitas* "you think," *cogitat* "he

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

Phase 1: Pretraining (+ feedback) 36 randomized trials	Phase 2: Sentence decoding (+ feedback) 36 (18 × 2) randomized trials	Phase 3: Reception testing (- feedback) 66 randomized trials	
	Present	Stimulus	Semidiem
Adverb pretraining group *Hodie* "today" *Heri* "yesterday"	*Hodie cogito* *Hodie cogitas* *Hodie cogitat* *Cogito hodie* *Cogitas hodie* *Cogitat hodie*	*Hodie* *Heri* *Cras* *Cogito/as/at* *Cogitavi/visti/vit* *Cogitabo/bis/bit* *Hodie cogito/as/at* *Hodie cogitavi/visti/vit* *Hodie cogitabo/bis/bit* *Heri cogito/as/at* *Heri cogitavi/visti/vit* *Heri cogitabo/bis/bit* *Cras cogito/as/at* *Cras cogitavi/visti/vit* *Cras cogitabo/bis/bit* *Cogito/as/at hodie* *Cogitavi/visti/vit hodie* *Cogitabo/bis/bit hodie* *Cogito/as/at heri* *Cogitavi/visti/vit heri* *Cogitabo/bis/bit heri* *Cogito/as/at cras* *Cogitavi/visti/vit cras* *Cogitabo/bis/bit cras*	3 1 5 3 1 5 3 2 4 2 1 3 4 3 5 3 2 4 2 1 3 4 3 5
Verb pretraining group *Cogito* "I think" *Cogitas* "you think" *Cogitat* "X thinks" *Cogitavi* "I thought" *Cogitavisti* "you thought" *Cogitavit* "X thought"	Past *Heri cogitavi* *Heri cogitavisti* *Heri cogitavit* *Cogitavi heri* *Cogitavisti heri* *Cogitavit heri*		
Control group No pretraining	Future *Cras cogitabo* *Cras cogitabis* *Cras cogitabit* *Cogitabo cras* *Cogitabis cras* *Cogitabit cras*		

Figure 1. The design of Phase 1, Phase 2, and Phase 3 of Experiment 1. The rating scale for Phase 3 ranged from 1 (*past*) to 5 (*future*).

thinks") verb inflections and learned that each corresponded to one of the two translations *X think(s)* or *X thought*, clicking the appropriate alternative with the mouse. A correct choice returned the feedback "correct," whereas an incorrect one returned, for example, "wrong—the meaning of *cogitavi* is *I thought*." The 36 trials in Phase 1 of verb pretraining thus comprised each of the three persons of present and past tense presented six times in random order.

Note that for both groups, the stimuli presented in Phase 1 involved only present and past temporal reference, no future. Participants in the no pretraining control group had no Phase 1. After these different beginnings, all three groups of participants underwent identical Phase 2, Phase 3, and Phase 4.

Phase 2. Participants were exposed to 18 sentences (*hodie cogito*/***as***/***at*** "today I/you/he think(s)," *cogito*/***as***/***at*** *hodie* "I/you/he think(s) today," *heri cogitavi*/***visti***/***vit*** "yesterday I/you/he thought," *cogitavi*/***visti***/***vit*** *heri* "I/you/he thought yesterday," *cras cogitabo*/***bis***/***bit*** "tomorrow I/you/he will think," *cogitabo*/***bis***/***bit*** *cras* "I/you/he will think tomorrow") that appropriately combined the adverb with a verb—half in adverb-verb word order and half in verb-adverb order—and had to

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

choose whether these sentences referred to the present, the past, or the future. Both word orders were used to counterbalance which cue was experienced first across sentences. Each of the 18 sentences was presented twice during this phase of the experiment. Again, participants were given feedback if incorrect.

Phase 3. As in competition model studies of cue use (MacWhinney, 1987), in the reception test of Phase 3, all single-word items (adverbs *hodie* "today," *heri* "yesterday," *cras* "tomorrow" and verbs *cogito* "I think," *cogitas* "you think," *cogitat* "he thinks," *cogitavi* "I thought," *cogitavisti* "you thought," *cogitavit* "he thought," *cogitabo* "I will think," *cogitabis* "you will think," *cogitabit* "he will think") were presented either alone or in all their possible combinations, once in each of two possible word orders, for a total of 66 test items: 12 single-word items + 54 (27 combinations × 2 word orders) two-word items. Both word orders were used to counterbalance which cue was experienced first across these test sentences. On each trial, participants were asked to judge the temporal reference of each string on a 5-point scale. The factorial crossing of these cues resulted in sensible strings that the learners had experienced before, equivalent to, for example, *today I think* and *tomorrow I will think*; words that they had not experienced separately before such as *cras* "tomorrow" and *cogitabo* "I will think"; and strange combinations equivalent to, for example, *yesterday I think*. The possible scale points were labeled 1 *past*, 2 *between past and present*, 3 *present*, 4 *between present and future*, and 5 *future*. Participants were asked to type the number that they thought best applied. The responses to the individual word items and to the trials in which the adverbial and inflectional cues conflict allow the assessment of the relative weight that learners put on these different cues to temporal reference. The average weight of the two cues within each string, which will be called the *semidiem* (this being a Latin experiment), is provided on the right side of the Phase 3 panel of Table 1. There was no feedback in Phase 3.

Phase 4. Finally, in the production test of Phase 4, participants were asked to translate from English to Latin by typing in the Latin equivalents of a selection of the various elements to which they had been exposed. Nine unique items were given twice, for a total of 18 items. Three adverbs were given in isolation (*yesterday*, *today*, *tomorrow*), three verbs were given in isolation (*X thought*, *X think(s)*, *X will think*), and three sentences were given with the tense of the verb and the temporal adverb matching (*yesterday X thought*, *today X think(s)*, *tomorrow X will think*). All of the target productions in this portion had been used in Phase 2 of the experiment.

The logic of the experiment follows that of Ellis and Sagarra (2010). In Phase 2, every sentence comprises two cues—an adverb and a

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

morphological inflection, which both cue the same temporal reference. Every participant experiences these cues together. Control participants only ever see them together. If participants pay equal attention to these two cues, then in Phase 3, their judgments should be equally affected by them, and in Phase 4, they should be equally good at producing adverbs and verbal inflections. If, however, participants pay more attention to the adverbial (or verbal) cues, then their judgments will be swayed from the semidiem toward the adverbial (or verbal morphological) cues in Phase 3, and they will better produce them in Phase 4.

Control group performance in Phase 3 and Phase 4 thus indicates how native English speakers naturally weigh these two cues—whether, for example, more salient lexical cues overshadow less obvious morphological ones. Performance in the adverb pretraining group assesses potential learned attention and blocking as a detrimental effect of prior learning of lexical cues upon later learning of inflectional cues. Performance in the verb pretraining group assesses potential blocking as a detrimental effect of prior learning of inflectional cues upon later learning of lexical cues.

Subsequent responses relating to past and present judgments could reflect specific prior-learned associations as proactive interference effects where, as in paired associate learning experiments, memory for association A-B (e.g., present-*cogito* "I think") is worse after prior learning of A-C (present-*hodie* "today") in comparison with a control condition involving prior learning of unrelated material D-E (Baddeley, 1976). However, the future reference sentences *cras cogitabo* "tomorrow I will think" and *cogitabo cras* "I will think tomorrow" are a special case, in that every participant, whether in the control, verb pretraining, or adverb pretraining group, only ever experiences these two cues together in Phase 2. Nobody has prior experience of these in Phase 1. They are thus a pure indicator of learned attention to adverbial or verbal cue type. Statistical analyses were performed using the statistics package R (R Development Core Team, 2008).

Results

Phase 2. By the second half of Phase 2, control group performance, despite a lack of prior training, was 75% correct compared to 97% for the adverb and 87% for the verb pretraining conditions.

Phase 3 Perception Data. Participants in the three groups differed in their cue use in Phase 3. Figure 2 illustrates the average group understanding of the temporal reference of each of the constructions of Phase 3 in terms of deviations from the semidiem averages given in Figure 1.

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

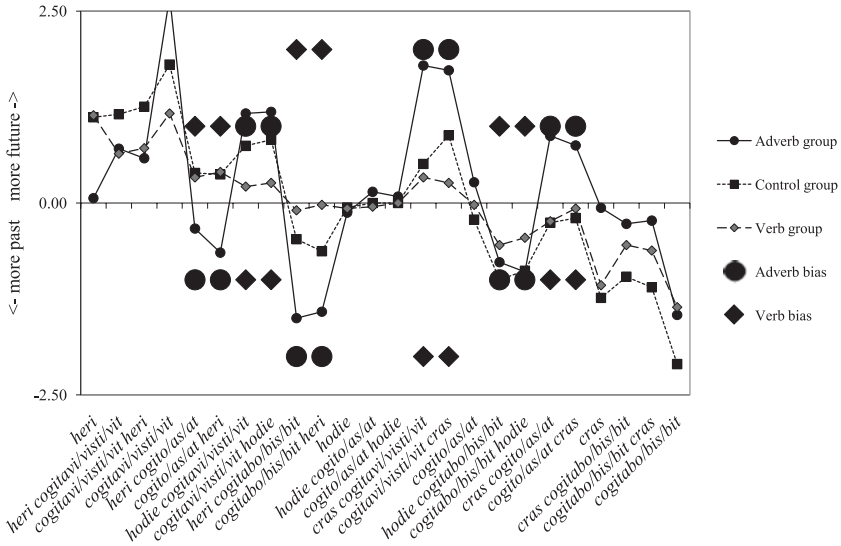


Figure 2. Mean deviation of Phase 3 temporal interpretations from semidiem average. Bias symbols mark the deviation of the adverbial cues (circles) and verb inflection cues (diamonds) when these cues conflict.

The strings are ordered from past on the left to future on the right. For each string where the two cues are in conflict, the large solid diamond shows the temporal information provided by the verb and the large solid circle shows the temporal information provided by the adverb. In the sentence *cogito heri* "I think yesterday," for example, with a verb reference of 3 and an adverb reference of 1, the respective deviations from the semidiem (2) are +1 and –1; for *cogitabo heri* "I will think yesterday," the verb reference is 5, the adverb reference is 1, and thus the respective deviations from the semidiem (3) are +2 and –2.

Participant ratings are similarly plotted as deviations from the semidiem. For example, consider a participant rating *heri cogitavi* "yesterday I thought." The semidiem in this case is 1: (1 [*yesterday*] + 1 [*I thought*])/2. Thus, a rating of 2 would be plotted as +1 (a one-unit deviation to the future: 2 [rating] – 1 [semidiem]); a rating of 1 would be plotted as 0 (zero deviation: 1 [rating] – 1 [semidiem]), and so on. As a second example, consider a participant rating *cras cogito* "tomorrow I think." The semidiem in this case is 4: (5 [*tomorrow*] + 3 [*I think*])/2. So a rating of 3 would thus be plotted as –1 (a one-unit deviation to the past: 3 [rating] – 4 [semidiem]; this is the score that someone who only interpreted *cogito* "I think" would get), a rating of 4 would be plotted as 0 (zero deviation: 4 [rating] – 4 [semidiem], indicating that the participant

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

was balancing the information from both cues), a rating of 1 would be plotted as –3 (a three-unit deviation to the past: 1 [rating] – 4 [semidiem]), and a rating of 5 would be plotted as +1 (a one-unit deviation to the future: 5 [rating] – 4 [semidiem]; this is the score that someone who only interpreted *cras* "tomorrow" would get), and so on.

Figure 2 plots these scores averaged across the participants within each of the three groups. Figure 2 shows that in the two-word strings of Phase 3, when the temporal information cued by the adverb and the verbal inflection differs, respondents overall tend to be more influenced by the adverb than the verb. However, the three groups are differentially sensitive to verb inflections: The adverb pretraining group is not at all sensitive to these cues, the control group is somewhat so—but still more attracted by the adverb—and the verb pretraining group is the most sensitive of the three groups to the verb cues.

These impressions are confirmed by three multiple regression analyses—one for each group—where the dependent variable is group mean temporal interpretation for each of the 54 two-word strings and the independent variables are the interpretations cued by the adverbial cue and by the verbal inflection. The differential cue use by each of the three groups, in standardized β coefficients, is shown in Table 2.

Participants in the control group, who had not been pretrained on either cue, were more reliant upon the adverbial cue ($\beta = 0.93$, $p < .001$) than the verbal cue ($\beta = 0.17$, $p < .001$); adverbial cues determined 86% of their ratings (0.93^2), whereas verbal morphology accounted for just 3%. The confidence intervals (CIs) of the two coefficients are nonoverlapping. This performance might reflect the relative salience, simplicity, and reliability of the adverbial cues compared to the verbal inflections.

Against this baseline, effects of prior exposure to cues can be measured. Participants in the adverb pretraining group were wholly reliant

Table 2. Regression analyses predicting mean temporal interpretation in Phase 3

Group	β	95% CI	Adjusted R^2	F
Adverb pretraining				
Adverb	0.99*	[0.96, 1.02]	0.99	2053*
Verb	–0.01	[–0.04, 0.02]		
Verb pretraining				
Adverb	0.76*	[0.69, 0.83]	0.94	413*
Verb	0.60*	[0.54, 0.67]		
Control				
Adverb	0.93*	[0.84, 1.02]	0.88	192*
Verb	0.17*	[0.07, 0.26]		

* $p < .001$.

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

upon the adverbial cue (β = 0.99, $p < .001$) and seemed to be unaffected by the verbal cue (β = –0.01, $p = .41$). Adverbial cues determined 99% of their ratings, the 95% CI spanning 1.0, whereas verbal morphology accounted for 0%, the 95% CI spanning 0.0. Participants in the verb pretraining group were reliant upon both the adverbial (β = 0.76, $p < .001$) and verbal (β = 0.60, $p < .001$) cues. Adverbial cues determined 58% of their ratings, whereas verbal morphology accounted for 37%.

Comparing the 95% CIs for these coefficients across groups, adverb pretraining and control participants were equally reliant upon the adverb cue and significantly more so than the verb pretraining participants. Verb pretraining participants were significantly more affected by the verbal cues than are the other two groups, with control participants, in turn, being significantly more sensitive to these cues than adverb pretraining participants. These impressions are confirmed using pooled variance CI estimates for the differences between these β coefficients (Edwards, 1984), which show (a) for adverb cue sensitivity, the adverb pretraining and control groups were more sensitive to adverb cues than was the verb pretraining group, $p < .01$; (b) for verb cue sensitivity, the verb pretraining group was most sensitive to these cues, followed by the control group, and then the adverb pretraining group, $p < .01$.

These differences in the relative amounts of variance explained by the adverb and verb cues on the mean group ratings over the test strings in Phase 3 are substantial. Nevertheless, as with all learning experiments, it is appropriate to ask whether the group performance means are truly reflective of the individuals within that group or whether they provide a central tendency that blurs individual within-group differences. To determine whether these patterns are reliable across individual group members, for each individual's responses in Phase 3, the degree to which their temporal rating on each construction correlated with the information provided by the verb cue and that provided separately by the adverbial cue was calculated. These Pearson's r correlations thus show the degree to which each participant was biased by each cue. Figure 3, which plots each individual in the space defined in this way, shows all of the adverb pretraining individuals to be heavily influenced by the adverb cue and not at all by the verbs. Most of the individuals in the control group were also predominantly influenced by the adverb cue, with only 4 group members showing any effect of verb cue greater than $r = 0.2$. The participants in the verb pretraining group are more scattered: 7 behave like members of the adverb group and use the adverb cue alone, but there are 9 group members who show sensitivity to the verb information at $r > 0.2$. Very few participants lie along the 45° diagonal, equally affected by these two cues, as the β weights for the verb group means might suggest. This finding is in line with others demonstrating that in the early stages of acquisition of learning about a

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

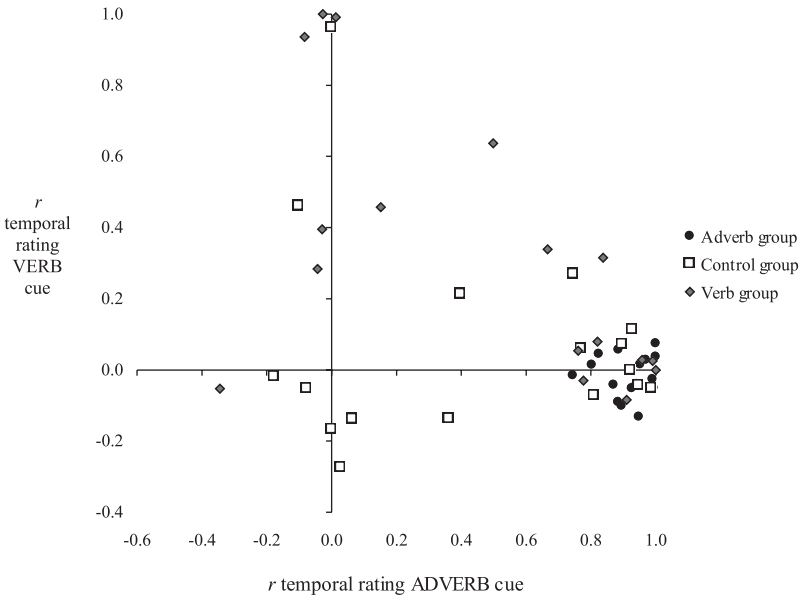


Figure 3. Sensitivity to adverbial and verbal inflectional cues to temporal reference in each participant.

problem space that comprises multiple cues to interpretation, participants typically focus upon one cue at a time, exploring its utility and only introducing others later, one-by-one, as they reduce error of estimation (Cheng & Holyoak, 1995; MacWhinney, 1987; Matessa & Anderson, 2000; McDonald, 1986). It also replicates the pattern found in Ellis and Sagarra (2010).

The group means of these correlations are shown in Figure 4. These group means of the individual correlations within each group are slightly different from the correlations of the group mean scores over the individuals because of the different orders of steps of calculation, but the patterns are substantially the same. The adverb pretraining group continues to be exclusively influenced by the adverb (M = 0.96, 95% CI [0.94, 0.98]) but not the verb (M = –0.02, 95% CI [–0.05, 0.015]). Control participants are more influenced by the adverb (M = 0.48, 95% CI [0.23, 0.73]) than the verb (M = 0.07, 95% CI [–0.13, 0.26]), whereas the verb pretraining group is influenced by both the adverb (M = 0.49, 95% CI [0.22, 0.76]) and verb (M = 0.38, 95% CI [0.17, 0.59]). An ANOVA shows this to be a highly significant interaction, $F(2, 47) = 8.33$, $p < .001$. Post hoc testing using Tukey's honestly significant differences (HSD) test showed that the adverb pretraining group differed from both the control and verb pretraining groups in adverb cue use ($p < .001$) but that these two groups did not differ significantly. The verb pretraining group differed significantly from the

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

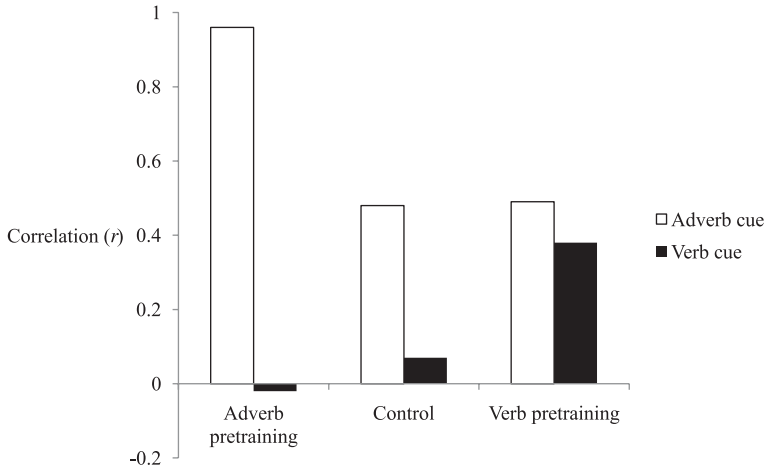


Figure 4. Group mean correlations between individual participants' Phase 3 sentence ratings and the information given by the corresponding adverb and verb cues.

adverb pretraining group in verb cue use ($p < .05$) but not from the control group ($p = .15$), and these latter two groups did not differ significantly.

Phase 4 Production Data. The production data parallel these patterns. Table 3 shows two typical individuals' responses from Phase 4 of the experiment, during which participants were asked to produce, for the first time, utterances in Latin that related to event time. The adverb pretraining participant provided the appropriate adverb on every trial where required—and often on trials where they were not—alongside a relatively unchanging central tendency of verb form, *congitavi* (an idiosyncratic blend of their own invention). The verb pretraining participant usually provided the appropriate tense inflections where required. When asked for a bare adverb, this participant did provide one—although there was some confusion over *hodie* "today" and *heri* "yesterday," which suggests that they had started to acquire the items in this category but without yet realizing their proper mappings to meaning. Phase 4 was less transparent for control group participants because they had not learned the translations in Phase 1, having started the experiment with Phase 2. Nevertheless, most control participants attempted this phase and, like language learners the world over, made a reasonable stab at generalizing from what they already knew.

To analyze the production data across individuals, each response was given a score of 0 to 1 for adverb or verb based on the following criteria: For the adverb, a score of 0 was given if no adverb was provided or the wrong adverb was provided, and a score of 1 was given if the adverb

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

was given and spelled correctly. Intermediate scores of .33 were given if the adverb was misspelled by two letters but remained a close match to the target (e.g., "codi" for *hodie* "today") and .66 if the correct adverb was given but misspelled by only one letter. For the verbs, a key factor in the scoring was the distinction between using a *b* in the ending for the future tense, a *v* in the verb ending for the past tense, and a lack of *b* or *v* for the present tense. A score of 0 was given if no verb was provided, and a score of 1 was given for the correct verb and correct spelling. Intermediate scores of .33 were given if a unique verb form was used to represent this tense—albeit one that did not make the correct distinction between *b* or *v*—or the correct future-past distinction was used (*b* vs. *v*) but the root was badly misspelled; a score of .66 was given for answers that included *cog* or *cogit* as the root and included correct use of the future-past distinction (*b* vs. *v*) but included some slight misspelling.

Each participant translated nine unique items (listed in Table 3) twice, for a total of 18 items. Among these 18 items, 12 required adverb production and 12 required verb production. Thus, each participant received two scores—one for adverb and one for verb—each a score out of 12. The group means of these production scores are shown in Figure 5. The adverb group participants were able to produce adverbs (M = 10.18, 95% CI [7.98, 11.39]) much more accurately than verbs (M = 1.08, 95% CI [0.02, 1.35]). The control group also produced more adverbs (M = 5.58, 95% CI [2.88, 7.47]) than verbs (M = 2.52, 95% CI [–0.03, 2.50]). In parallel to their perception data, the verb group showed knowledge of both adverbs (M = 7.27, 95% CI [5.13, 8.63]) and verbs (M = 5.82, 95% CI [1.74, 6.02]).

An ANOVA shows this to be a highly significant interaction, $F(2, 47)$ = 14.82, $p < .001$. Post hoc testing using Tukey's HSD test showed that the adverb pretraining group differed from the control group in adverb cue provision ($p < .01$) but not from the verb pretraining group ($p = .10$) and that these latter two groups did not differ significantly. The verb pretraining group was superior to both the adverb pretraining ($p < .001$) and control ($p < .05$) groups in verb cue provision, with no significant difference between these latter two groups.

These results mirror the perception data; the adverb pretraining group and the control group attended to adverb cues and were able to produce this information accurately to a relatively high proficiency; conversely, the production of verb information was low, even when no pretraining had occurred as in the control group. The verb group was superior at both the perception and production of verbal cues to temporal reference.

Cues or Content? A key element of the design of the experiment is that all participants had equivalent exposure to the future items *cogitabo*/*bis*/*bit cras* "I/you/he will think tomorrow" and *cras cogitabo*/*bis*/*bit*

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

Table 3. Representative responses from Phase 4 production testing

Stimulus	Correct response	Typical adverb group response	Typical verb group response
X thought	*Cogitavi / cogitavisti/ cogitavit*	*X congitavi*	*Cogitavi*
X think(s)	*Cogito / cogitas / cogitat*	*Congitavi hodie*	*Cogitat*
X will think	*Cogitabo / cogitabis / cogitabit*	*Congitavi cras*	*Cogitabo*
Yesterday	*Heri*	*Heri*	*Hodie*
Today	*Hodie*	*Hodie*	*Heri*
Tomorrow	*Cras*	*Cras*	*Curas*
Yesterday X thought	*Heri cogitavi / cogitavisti / cogitavit*	*Heri X congitavi*	*Hodie cogitavi*
Today X think(s)	*Hodie cogito / cogitas/ cogitat*	*Congitavi hodie*	*cogitat*
Tomorrow X will think	*Cras cogitabo / cogitabis / cogitabit*	*Cras congitavi*	*Cras cogitabo*
X thought	*Cogitavi / cogitavisti/ cogitavit*	*X congitavi*	*Cogitavi*
X think(s)	*Cogito / cogitas / cogitat*	*X congitavi*	*Cogitat*
X will think	*Cogitabo / cogitabis / cogitabit*	*X congitavas?*	*Cogitabo*
Yesterday	*Heri*	*Heri*	*Heri*
Today	*Hodie*	*Hodie*	*Hodie*
Tomorrow	*Cras*	*Cras*	*Cras*
Yesterday X thought	*Heri cogitavi / cogitavisti / cogitavit*	*Heri X congitavi*	*Hodie cogivisti*
Today X think(s)	*Hodie cogito / cogitas/ cogitat*	*Hodie X congitavi*	*Heri cogitat*
Tomorrow X will think	*Cras cogitabo / cogitabis / cogitabit*	*Cras congitavas*	*Cras cogita*

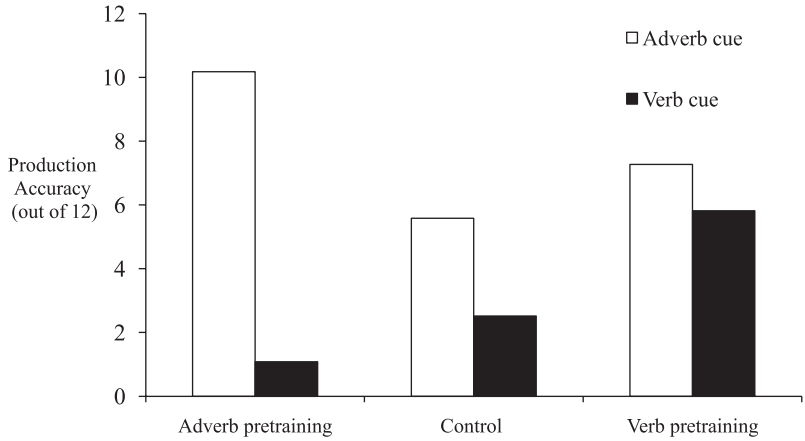


Figure 5. Group mean production scores for adverb and verb cues.

"tomorrow I/you/he will think" in Phase 2 and that no one had any experience with these items or any reference to future temporality beforehand. Every participant in the experiment encountered each of these six utterances twice in Phase 2 and was shown that they relate to the future. Analysis of these items therefore allows for the identification of pure effects of blocking devoid of prior training on particular content. Performance on these strings must reflect attentional biases to particular dimensions of cue (adverb vs. verbal inflection) rather than to particular words or tenses because neither these forms nor their functions have been encountered before Phase 2.

Inspection of the right side of Figure 2 suggests that the adverb and verb pretraining groups are as unalike and dissociated in their performance on *cras* "tomorrow" and *cogitabo*/*bis*/*bit* "I/you/he will think" items referring to the future as they are on the past and present reference items in Phase 3.

Figure 6 isolates these future items for the adverb pretraining and verb pretraining groups. The left panel shows the two groups' temporality ratings for the adverb *cras* "tomorrow" (adverb pretraining, M = 4.94, 95% CI [4.80, 5.07]; verb pretraining, M = 4.06, 95% CI [3.44, 4.67]) and the verb inflection *cogitabo* "I will think" (adverb pretraining, M = 3.54, 95% CI [3.11, 3.97]; verb pretraining, M = 3.61, 95% CI [3.13, 4.08]) when they are experienced on their own as individual words in Phase 3. Both of these items, if fully acquired, should receive a future rating of 5. However, the adverb pretraining group has learned more from the two-word utterances experienced in Phase 2 about *cras* "tomorrow," whereas the verb pretraining group has learned more from the same exposure to the same utterances about *cogitabo* "I will think." A two-factor ANOVA (2 groups × 2 cues with participants nested within groups)

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

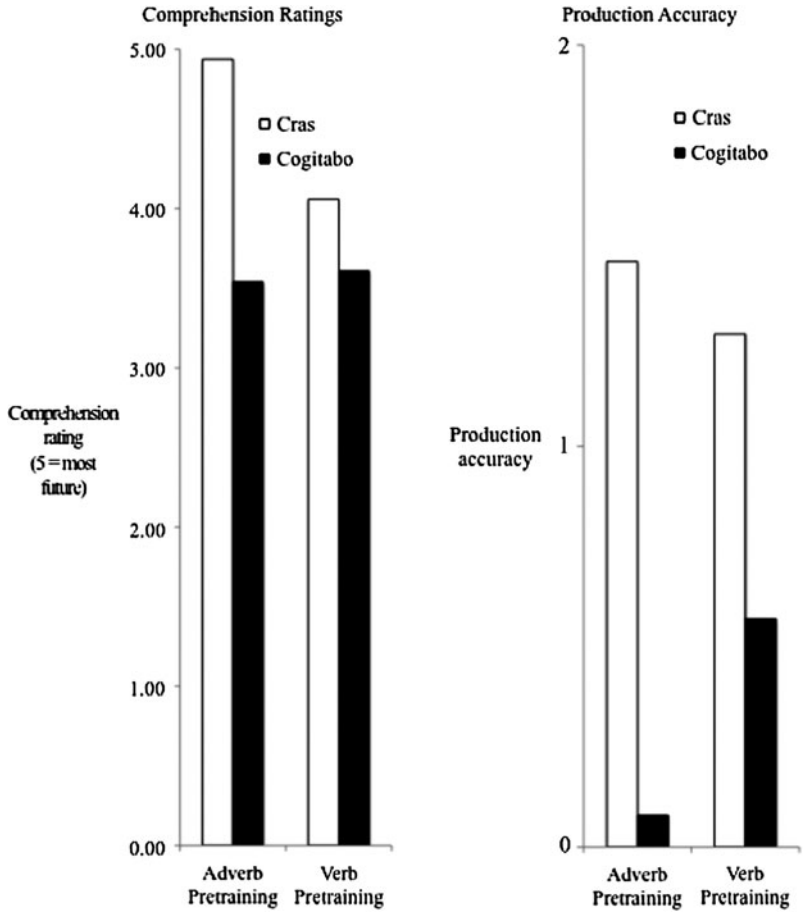


Figure 6. Adverb pretraining and verb pretraining group comprehension rating (left) and production (right) of *cras* "tomorrow" and *cogitabo* "I will think."

shows a main effect of group with the adverb pretraining group more accurate overall, $F(1, 31) = 4.99$, $p < .05$, a main effect of cue with higher overall accuracy on the adverb cue, $F(1, 31) = 13.93$, $p < .001$, and a significant two-way interaction, $F(1, 31) = 3.76$, $p < .05$ (one tailed).

These items can similarly be isolated in production. In Phase 4, the participants were asked for the Latin translations of, on separate trials, *tomorrow*, *I will think*, and *tomorrow I will think*. The right panel of Figure 6 shows the summed accuracy (out of a possible maximum 2.0) of the two groups on the adverb *cras* "tomorrow" (adverb pretraining, $M = 1.46$, 95% CI [1.02, 1.90]; verb pretraining, $M = 1.28$, 95% CI [0.86, 1.70]) and the verb inflection *cogitabo* "I will think" (adverb pretraining, $M = 0.08$, 95% CI [–0.02, 0.19]; verb pretraining, $M = 0.57$, 95% CI [0.24, 0.89]).

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

Although both groups seem to have picked up something about *cras* from the two-word utterances experienced in Phase 2, the verb pretraining group has learned more from that same exposure to these utterances about *cogitabo*. A two-factor ANOVA (2 groups × 2 cues with participants nested within groups) shows a nonsignificant main effect of group, $F(1, 31) = 0.96$, $p = .33$, a main effect of cue with higher overall accuracy on the adverb, $F(1, 31) = 35.89$, $p < .001$, and a significant two-way interaction, $F(1, 31) = 3.60$, $p < .05$ (one tailed).

Discussion

The results of Experiment 1 demonstrate that salient adverbs were better learned than less salient verb inflections, that early experience of adverbial cues led to subsequent reliance upon these rather than verbal inflections, and that making verbal inflections salient in pretraining allowed participants to be sensitive to these cues in the processing of subsequent utterances that contained both adverbial and morphological cues to temporal reference. The fact that these patterns were found with the constructions relating to future temporality, which had not been experienced before Phase 2, points to the conclusion that these results reflect effects of learned attention to cue dimensions in the early acquisition of language.

The results therefore replicate earlier findings but also extend them. In Experiment 1 of Ellis and Sagarra (2010), participants in the English L1 control condition (which involved just one verb form per tense) showed a differential cue use to temporal interpretation, in standardized β coefficients, of 0.60 Adverb + 0.72 Verb. In the current experiment, with three verb forms per tense, the English L1 control group exhibited cue use of 0.93 Adverb + 0.17 Verb. This difference might reflect the relative salience, simplicity, and reliability of adverbial cues compared to verbal inflections and the fact that adult language learners have prior knowledge of the use of adverbial temporal references from their L1. This issue will be revisited in the general discussion.

These findings demonstrate that for linguistic constructions too—as in the case of associative learning of other cue-outcome interpretations in medical diagnosis or in stock market prediction (Kruschke & Blair, 2000; Shanks, 1995)—prior learning of one cue affects learner attention so that they subsequently attend to that cue more than to others, however reliable these are as predictors in their own right. This experiment shows these learned attention effects in the short-term of an experiment that manipulated instructional sequence. The next experiment concerns whether learned attention also occurs in the longer term.

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

EXPERIMENT 2

Usage-based, constructionist views of language acquisition (Barlow & Kemmer, 2000; Collins & Ellis, 2009; Ellis, 1998, 2003; Ellis & Cadierno, 2009) hold that language experiences accumulate over time and that from their integration there emerges the processes, representations, and attentional biases that constitute the language system. Experience of how the L1 maps on to experience colors expectations of a second language (L2), resulting in crosslinguistic transfer effects (MacWhinney, 1997; Odlin, 1989; Robinson & Ellis, 2008). At least to some extent, limited adult language attainment is grounded in L1 entrenchment, learned attention, and transfer, rather than in age or biology per se (Ellis, 2006c).

Experiment 2 of Ellis and Sagarra (2010) showed that native speakers of Chinese (whose L1 does not exhibit verb tense morphology) learning under the control conditions of Experiment 1, where there was just one verb form per tense, were less sensitive to morphological cues to tense and more reliant on adverbial cues than native English speakers. In Chinese languages, there is no morphology, free or bound, that corresponds to tense: "a second factor in the lack of inflectional morphology is that gender, plurality and tense are either indicated by lexical choice or not indicated at all" (Li & Thompson, 1987, p. 825). Mandarin Chinese does have aspectual markers, including the perfective marker *–le*, an inflection that cannot be separated from the verb, which indicates a recently completed event that has present relevance. This fact coupled with the absence of verbal inflectional morphology renders its characterization as an exclusively aspect language (Lin, 2006; Xiao & McEnery, 2004). Instead of verbal tense inflections, speakers make heavy use of temporal adverbials—both adverbs and prepositional phases—to encode temporal meanings. Ellis and Sagarra therefore concluded that Chinese speakers' L1 experience had summed to long-term attentional biases toward these types of cue, with consequent blocking of verbal inflectional cues in the Latin learning experiment. Again, this is a result that warrants replication and extension to both a richer target verb inflectional system as in Experiment 1 and to L1s that make more use of tense morphology than English does.

In contrast to Chinese, many languages present a much richer verbal morphology than English, containing overt marking for number, person, and gender agreement as well as tense and mood. Two typical languages that inflect in a similar fashion to Latin are Spanish and Russian. All three of these languages evidence verbal inflection that agrees with the subject in number and person through suffixation. Russian makes a binary morphological distinction between past and nonpast (i.e., present-future). In the nonpast, verbs agree with their subject in person and number; in the past, they agree in gender and number (Comrie, 1987). Spanish distinguishes among the future,

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

the present, the preterit, and the imperfect, with each paradigm consisting of six forms that represent three grammatical persons in both singular and plural (Green, 1987). These characteristics are, of course, not limited to these three languages; most (if not all) languages in the Romance and Slavonic language families have rich verbal morphology with respect to these same features. Spanish and Russian, like Latin, are both also pro-drop or null-subject languages in which pronominal syntactic subjects are usually not present in the sentence (although they can be overtly pronounced in certain pragmatic contexts, e.g., to communicate contrastive stress). Because of this omission, the person and number information of the subject is encoded exclusively in the verbal morphology. Of course, in addition to the verbal inflectional ending that encodes tense, the language can make use of temporal adverbs, pragmatics, and other mechanisms to reinforce the interpretation of semantic tense outside the verbal domain. These aspects of Spanish and Russian lead to the expectation, therefore, that speakers of these L1s would have learned attentional biases toward verbal morphological cues and that these biases would transfer to FL learning. This experiment therefore investigates this issue.

It also addresses two further questions. The first is whether these long-term learned attention effects are a matter of degree. Are learners of an inflection-light L1 (like English) more sensitive to verbal morphology than inflection-free-L1 learners but less sensitive than inflection-rich-L1 learners? The second is whether there are separable effects of attention to cue dimension and attention to cue content. Although Russian and Spanish are both languages that make rich use of verbal morphology, Spanish is more similar to Latin in the way it cues time. Russian verbal morphology is synthetic, with more than one meaning mapping to a single morpheme. Spanish, a Romance language, derives via Latin from the Italic branch of Indo-European. Thus, it bears similarity to Latin in both its use and manner of verbal morphology, and although its overall degree of verbal morphology has declined somewhat from Latin levels, the most frequently occurring forms of the verb remain highly synthetic in structure. In its manner of inflection, it distinguishes among the future, the present, the preterit, and the imperfect, with each paradigm consisting of six forms that represent three grammatical persons in both singular and plural, largely with each meaning mapping to different principal parts or roots (Green, 1987). Verbs with Spanish infinitives in the *–ar* group derive directly from Latin, and there are, thus, clear similarities in the Latin and Spanish paradigms (compare Latin and Spanish present *cogito*/*as*/*at* "I/you/he think(s)" and *tomo*/*as*/*a* "I/you/he take(s)," past *cogitavi*/*visti*/*vit* "I/you/he thought" and *tomé*/*aste*/*ó* "I/you/he took," and future *cogitabo*/*bis*/*bit* "I/you/he will think" and *tomaré*/*ás*/*á* "I/you/he will take").

Experiment 2 therefore compares the learning of four L1 groups (inflection-free Chinese, inflection-light English, and inflection-rich

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

Russian and Spanish—but with Spanish closer to Latin than is Russian in its means and content of tense inflections) for their learning of verbal morphology under the control conditions of Experiment 1, where each tense is represented in three persons.

Participants

Participants were 63 adults who were either students at a major university in the United States (n = 46) or volunteers from the local community (n = 17). They were placed in one of four groups based on their L1: native Spanish speakers (n = 15), native Russian speakers (n = 17), monolingual English speakers (n = 17), and native speakers of languages that do not evidence inflectional morphology (n = 14; 12 Chinese, 1 Malay, 1 Indonesian). The Spanish group consisted of 9 males and 7 females, age range 19–40 years (mean: 23.9; median: 21.0). The Russian group consisted of 10 males and 7 females, age range 19–60 years (mean: 23.8; median: 20.0); as can be seen from the measures of central tendency, although the range here appears large, there was just one 60-year-old, and the age distributions across the language groups were otherwise similar). The English group consisted of the same participants as the no pretraining control group of Experiment 1.

All participants in the first, second, and third groups were bilingual with advanced English language proficiency sufficient to allow their study at the university through the medium of English. Their performance in this experiment therefore approximates third-language acquisition. None of the participants were natively bilingual—they were selected providing their L2 learning had occurred postpuberty. All were volunteers and were paid $10 for their participation in the experiment.

Procedure

The participants partook in an exact replication of the no pretraining control group of Experiment 1. It comprised Phase 2 sentence decoding, Phase 3 reception testing, and Phase 4 production testing.

Results

The performance of the four L1 groups in terms of deviation from semidiem judgments in Phase 3 is shown in Figure 7. As in Experiment 1,

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

respondents tend to be more influenced by the adverb than the verb. However, the degree to which this is so depends on L1 background. In Figure 7, the Chinese group's performance lies to the adverb side of the English group's line, tracking the information given by the adverbial cue more than by that from the verbal morphology, whereas the performance of the Spanish and Russian groups lies on the other side, tracking more the information given by the verbal inflections. Sensitivity to these cues is, therefore, a matter of degree.

These impressions are confirmed by four multiple regression analyses, one for each group, in which the dependent variable is group mean temporal interpretation for each of the 54 two-word strings and the independent variables are the interpretations cued by the adverbial cue and by the verbal inflection. The differential cue use by each of the four groups, in standardized β coefficients, is shown in Table 4. Participants in the Chinese group were highly reliant upon the adverbial cue ($\beta = 0.95$, $p < .001$), the 95% CI spanning 1.0, but were unaffected by the verbal cue ($\beta = -0.02$, $p = .65$), the 95% CI spanning 0. Participants in the English and Russian groups were broadly similar in that they, too, were heavily influenced by the adverbial cues ($\beta = 0.93$, $p < .001$, and $\beta = 0.91$, $p < .001$, respectively), again their CIs spanning 1.0. However, both of these groups were sensitive to the verbal morphology ($\beta = 0.17$, $p < .001$, and $\beta = 0.22$, $p < .001$, respectively). Participants in the Spanish group were rather less influenced by the adverbial cues

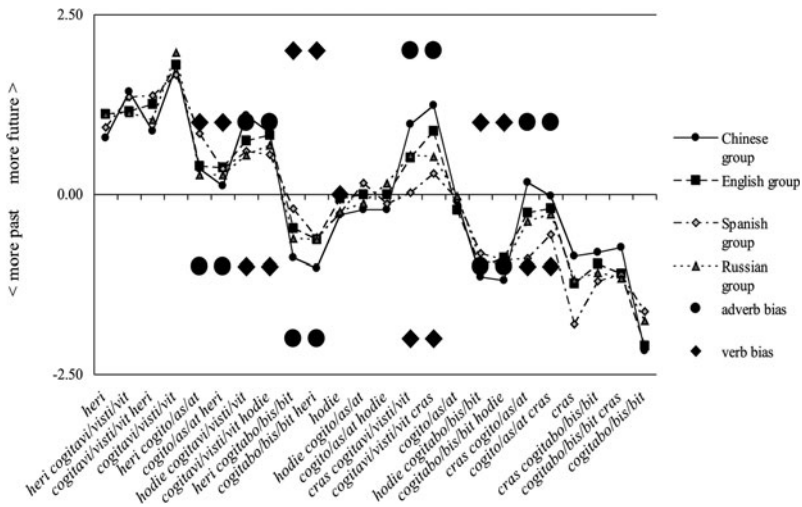


Figure 7. Mean deviations of Phase 3 temporal interpretations from semidiem average. Bias symbols mark the deviation of the adverbial cues (circles) and verb inflection cues (diamonds) when these cues conflict.

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

Table 4. Regression analyses predicting mean temporal interpretation in Phase 3 of Experiment 2

L1 Group	β	95% CI	Adjusted R^2	F
Chinese				
Adverb	0.95*	[0.86, 1.04]	0.89	221*
Verb	–0.02	[–0.11, 0.07]		
English				
Adverb	0.93*	[0.84, 1.02]	0.88	192*
Verb	0.17*	[0.07, 0.26]		
Russian				
Adverb	0.91*	[0.82, 1.01]	0.88	196*
Verb	0.22*	[0.13, 0.32]		
Spanish				
Adverb	0.75*	[0.60, 0.89]	0.71	66*
Verb	0.41*	[0.26, 0.55]		

* $p < .001$.

(β = 0.75, $p < .001$) and rather more the verbal inflections (β = 0.41, $p < .001$). Pooled variance CI estimates for the differences between these β coefficients (Edwards, 1984) showed the following differences for sensitivity (where < indicates less sensitivity and ≈ indicates relatively equivalent sensitivity) at $p < .05$ to the adverb cues: Spanish < Chinese, Spanish < English; Chinese ≈ English ≈ Russian; and to the verb cues: Chinese < English, Chinese < Russian, Chinese < Spanish; English < Spanish, Russian < Spanish.

As in Experiment 1, each L1 participant's temporal rating responses for the strings in Phase 3 were correlated with the information provided by the verb cue and that separately provided by the adverbial cue to determine the degree to which each participant was biased by each cue. Participant's production of the adverbs and verbs in Phase 4 was also scored as in Experiment 1. Figure 8 plots the group means for the rating and production data together, and again these illustrate a trend whereby sensitivity to these cues is determined by the degree to which the L1 makes use of inflectional morphology to mark tense.

However, as can be seen in the descriptive data for these measures in Table 5, there was considerable within-group variability. A two-factor ANOVA (4 groups × 2 cues with participants nested within groups) on the Phase 3 data demonstrated a nonsignificant main effect of group, $F(3, 59) = 0.25$, $p = .86$, a significant main effect of cue with more reliance, overall, on the adverb, $F(1, 59) = 24.12$, $p < .001$, and a nonsignificant Group × Cue interaction, $F(3, 59) = 1.56$, $p = .21$. The same analysis for the production data likewise resulted in a nonsignificant main effect of Group, $F(3, 59) = 0.24$, $p = .87$, a significant main effect of Cue with greater

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

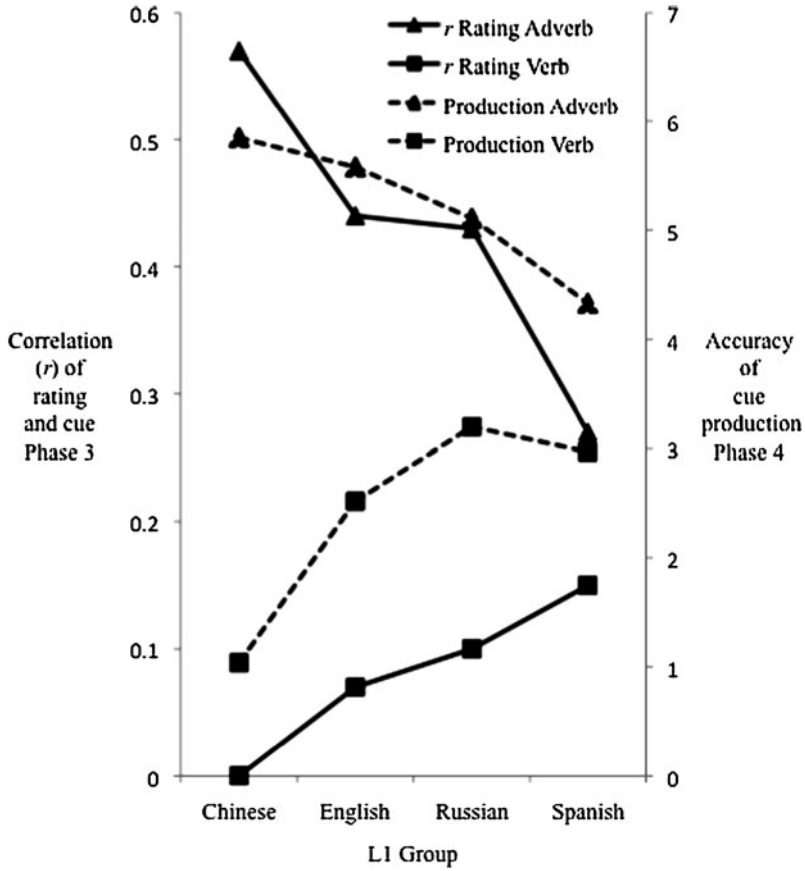


Figure 8. Means scores of the four L1 groups in Experiment 2 as they are affected by adverbial and verbal inflectional cues to temporal reference in comprehension in Phase 3 (left axis) and as they produce them in Phase 4 (right axis).

accuracy overall on the adverb cue, $F(1, 59) = 19.91$, $p < .001$, and a nonsignificant Group × Cue interaction, $F(3, 59) = 1.43$, $p = .24$.

This pattern of results is clarified by plotting in Figure 9 the position of each participant as defined by the correlation of their ratings of the sentences in Phase 3 with the information separately given by the verb and adverb cues. It can be seen that the majority of points cluster around the axes: Participants seem to be picking up on either one cue or the other and show different levels of mastery of this cue. There are no participants in the Chinese group who show substantial sensitivity to the verb cue. There are more participants in the Spanish and Russian L1 groups affected by the verb cue—but by no means everybody in these groups; there are still many who rely on the adverb. At these early

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

Table 5. Mean participant performance in Experiment 2

	Correlation of rating and cue (Phase 3)		Accuracy of cue production (Phase 4)	
L1 Group	Mean	95% CI	Mean	95% CI
Chinese				
Adverb	0.57	[0.30, 0.84]	5.85	[3.31, 8.39]
Verb	0.00	[–0.11, 0.11]	1.04	[0.21, 1.87]
English				
Adverb	0.44	[0.21, 0.67]	5.58	[3.26, 7.90]
Verb	0.07	[–0.08 0.22]	2.52	[1.15, 3.89]
Russian				
Adverb	0.43	[0.21, 0.65]	5.11	[2.81, 7.41]
Verb	0.10	[–0.06, 0.25]	3.20	[1.58, 4.82]
Spanish				
Adverb	0.27	[0.02, 0.52]	4.33	[1.88, 6.78]
Verb	0.15	[–0.03, 0.33]	2.97	[1.63, 4.30]

stages of learning, therefore, as in Experiment 1, participants seem to be picking up on one cue or another. The cue that they begin with is probabilistically determined by the degree to which the L1 uses inflectional morphology, but this is attenuated by a general bias toward the simpler adverbial cue.

A further question is related to whether there might be transfer of particular content—whether the similarity of inflectional system between Spanish and Latin might afford Spanish L1 participants a greater advantage over Russian L1 participants. Two-factor ANOVAs (2 groups × 2 cues, with participants nested within groups) demonstrated that the Group × Cue interaction failed to reach significance both for the rating data, $F(1, 30) = 0.93$, $p = .34$, and for the production data, $F(1, 30) = 0.09$, $p = .77$.

Discussion

The results of Experiment 2 demonstrate a pattern whereby, when exposed to utterances that contain both cues to temporality, respondents tend to be more influenced by the adverb than the verb. However, L1 speakers of inflection-light Chinese make relatively more use of adverbial cues in their processing, whereas L1 speakers of inflection-rich languages like Spanish and Russian are more sensitive to verbal inflectional cues. These findings confirm a long-term influence of attention to language, a cue processing bias that comes from a lifetime of prior L1

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

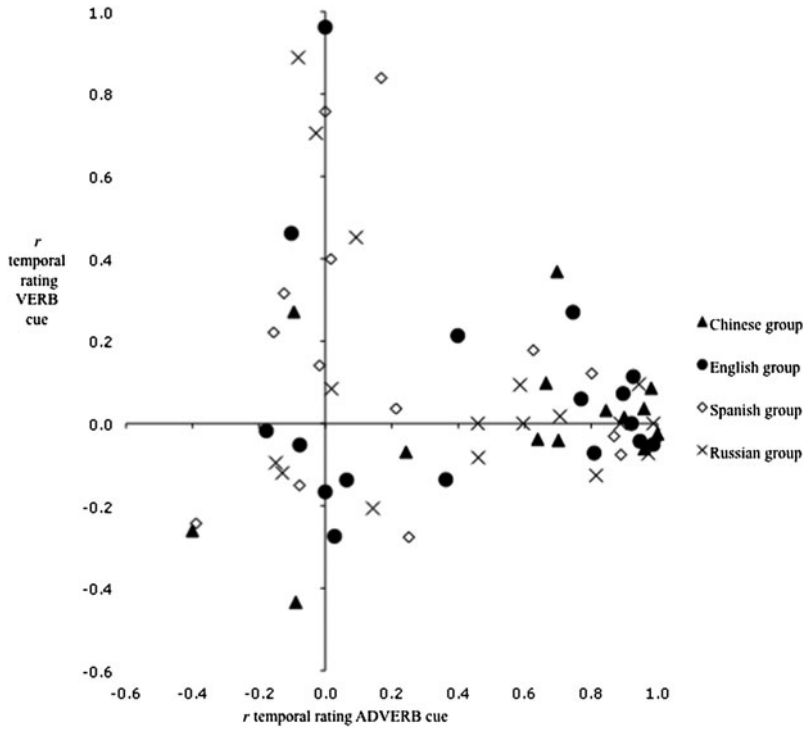


Figure 9. Sensitivity to adverbial and verbal inflectional cues to temporal reference in each participant.

usage. It is especially compelling, in that the Chinese L1 participants had been exposed to a subsequent L2 prior to the Latin learning experiment, the English in which they had become quite proficient and which, as a L2 learning experience, must have brought to their awareness the potential productivity of inflectional cues in tense marking. It would be interesting to compare the performance of this group with that of monolingual Chinese speakers who have not been potentially sensitized in this way.

As in Experiment 1, the current results replicate earlier findings but also extend them. Ellis and Sagarra's (2010) Experiment 2 showed that native speakers of Chinese (whose L1 does not exhibit verb tense morphology) learning under the control conditions of Experiment 1—where there was just one verb form per tense—showed a differential cue use to temporal interpretation, in standardized β coefficients, of 0.91 Adverb + 0.29 Verb. In the current experiment, with three verb forms per tense, the Chinese L1 control group cue use was 0.95 Adverb – 0.02 Verb. This shift in reliance away from the inflectional cues as the paradigm is complexified parallels that found in the English controls in Experiment 1. This difference too might reflect the relative salience, simplicity, and

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

reliability of adverbial cues compared to verbal inflections and the fact that adult language learners have prior knowledge of the use of adverbial temporal references from their L1—in the case of Chinese L1 speakers, exclusively so.

Together, these data point more to a cline of long-term learned attention to verbal morphology rather than a binary distinction. In both the comprehension and the production data, the Chinese group stood at one extreme, the Spanish the other, and the English group lay in between, with Russian group participants more similar to the English group on some measures and more similar to the Spanish group on others. There also seemed to be an effect of L1-FL relatedness whereby native Spanish speakers learning cognate Latin were more likely than native Russian speakers to attend to verbal morphology: For example, in the Phase 3 multiple regressions, Russian L1 speakers were less sensitive to morphological cues and more sensitive to the adverbs.

However, as these learners are considered individually, it becomes clear that these biases are stochastic. There are few individuals who are sensitive to both cues at these initial stages of learning; instead, participants pick up on one cue or another. The cue that they begin with is probabilistically determined by the degree to which their L1 has sensitized them to inflectional morphology, and they subsequently explore this cue dimension to varying degrees of achievement, a pattern also found in the other experiments of the current study and in Ellis and Sagarra (2010). Research looking at longer term learning from problem spaces containing multiple cue dimensions suggests that participants typically focus upon one cue at a time, exploring its utility and only introducing others later, one by one, as they reduce error of estimation (Cheng & Holyoak, 1995; MacWhinney, 1987; Matessa & Anderson, 2000; McDonald, 1986).

SEPARABLE EFFECTS OF ENTRENCHMENT AND BLOCKING: A META-ANALYSIS

Learned attention effects can be positive or negative. Experience can lead people to pay more attention to cues known in the past, to process them more fluently or automatically, to be more open to particular evidence, and to be positively disposed to arguments (positive). It can also lead people to ignore new, foreign cues, to be blind to information, and to be prejudiced against beliefs or explanations (negative). Although these are often related, they are logically separable effects. In Experiment 1, for example, adverb pretraining might cause participants to pay more attention to adverbs in future language processing, but their future sensitivity to other morphological cues could be untouched by this history. Positive effects of frequency of exposure are those of entrenchment

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

(see Ellis, 2002, for a review of frequency effects). Alternatively, increased sensitivity to adverb cues might be accompanied by a reduced sensitivity to morphological cues—that is, blocking as discussed in Ellis's (2006c) review of learned attention. Increased reliance on one cue, means, or method tends to diminish reliance on others but not necessarily so. Not only are these logically separable effects, but they are empirically separable as well. In Experiment 1, the precise history of Phase 1 preexposure of the adverb pretraining group to adverb cues is known, as is that of the verb pretraining group to verb cues. These two groups had the same subsequent exposure to adverb and verb cues in Phase 2, and this was equal to that of the control participants. Entrenchment effects are therefore indexed by increased sensitivity of the adverb group to adverbs and by increased sensitivity of the verb pretraining participants to verb cues, relative to control participants. Blocking effects are indexed by decreased sensitivity of adverb pretraining participants to verb cues and by decreased sensitivity of verb pretraining participants to adverb cues, relative to controls.

Tables 6 and 7 gather these data as effect sizes for entrenchment and blocking respectively across the various indexes of sensitivity analyzed in Experiment 1 of Ellis and Sagarra (2010) and the current study (βs from the group multiple regressions of cue use across the different test strings of Phase 3, means of individual participant *r*s indicating cue sensitivity in Phase 3, production accuracy in Phase 4). These measures do not obligate cue competition; they are independent, permitting equivalent accuracy on both cues up to maximum of 1.0 for β and *r* and 12.0 for production accuracy (in Ellis & Sagarra, 2010, this was expressed as a proportion). These different measures are converted into effect sizes to allow their comparability in a meta-analysis across measures and experiments. The pooled standard deviation of control and comparison groups was used to calculate the effect sizes in the meta-analyses.

A rule of thumb for judging effect sizes is as follows: $d \approx 0.2$ indicates a small effect; $d \approx 0.25$ might reflect educational significance (i.e., something was learned); $d \approx 0.5$ represents a moderate effect that is practically or clinically significant (i.e., something really changed); $d \geq 0.8$ indicates a large effect (Cohen, 1992; Thalheimer & Cook, 2002). Across measures and experiments, the average effect size for entrenchment is 1.23, a large effect, 95% CI [0.99, 1.47]. Ten of the 12 effect sizes for entrenchment in Table 6 are greater than 0.80. The average effect size for blocking is 0.52, a moderate effect, 95% CI [0.17, 0.90]. There does seem to be some identifiable variability across experiment and cue. Figure 10 plots those effect sizes broken down by experiment, cue, and entrenchment or blocking. Overall, entrenchment is a stronger effect than blocking. Effects are larger in Ellis and Sagarra (2010), in which the adverb and verb cues were of approximately the same salience than in the current study, in which the verb cues are less obvious than the adverbs. The salient adverbs in

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

Table 6. Evidence for entrenchment from Ellis and Sagarra (2010) and this replication study

Comparison	Comparison group			Control group			Effect size
	M	*SD*	*n*	*M*	*SD*	*n*	*d*
Experiment 1 of Ellis and Sagarra (2010)							
Multiple regression: Cue use in comprehension (β)							
Adverb pretraining-Control (adverb use)	0.97	0.13	36	0.59	0.39	36	1.31
Verb pretraining-Control (verb use)	0.97	0.20	36	0.72	0.39	36	0.83
Group mean correlation: Cue use in comprehension (*r*)							
Adverb pretraining-Control (adverb use)	0.78	0.26	19	0.27	0.41	16	1.54
Verb pretraining-Control (verb use)	0.80	0.13	17	0.42	0.38	16	1.37
Cue production (*M*)							
Adverb pretraining-Control (adverb use)	0.90	0.19	19	0.45	0.31	16	1.78
Verb pretraining-Control (verb use)	0.77	0.25	17	0.35	0.23	16	1.71
Experiment 1 of the Current Study							
Multiple regression: Cue use in comprehension (β)							
Adverb pretraining-Control (adverb use)	0.99	0.09	54	0.93	0.28	36	0.26
Verb pretraining-Control (verb use)	0.60	0.20	54	0.17	0.28	36	1.45
Group mean correlation: Cue use in comprehension (*r*)							
Adverb pretraining-Control (adverb use)	0.96	0.03	16	0.48	0.49	17	1.36
Verb pretraining-Control (verb use)	0.38	0.41	17	0.07	0.39	17	0.79
Cue production (*M*)							
Adverb pretraining-Control (adverb use)	10.18	2.72	16	5.58	4.51	17	1.23
Verb pretraining-Control (verb use)	5.82	3.31	17	2.52	2.66	17	1.10

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

Table 7. Evidence for blocking from Ellis & Sagarra (2010) and this replication study

Comparison	Comparison group			Control group			Effect size *d*
	M	*SD*	*n*	*M*	*SD*	*n*	
Experiment 1 of Ellis and Sagarra (2010)							
Multiple regression: Cue use in comprehension (β)							
Adverb pretraining-Control (verb use)	0.23	0.12	36	0.72	0.39	36	–1.70
Verb pretraining-Control (adverb use)	0.12	0.20	36	0.59	0.39	36	–1.52
Group mean correlation: Cue use in comprehension (*r*)							
Adverb pretraining-Control (verb use)	0.23	0.35	19	0.42	0.38	16	–0.53
Verb pretraining-Control (adverb use)	0.13	0.20	17	0.27	0.41	16	–0.43
Cue production (*M*)							
Adverb pretraining-Control (verb use)	0.41	0.33	19	0.35	0.23	16	0.21
Verb pretraining-Control (adverb use)	0.33	0.30	17	0.45	0.31	16	–0.37
Experiment 1 of the Current Study							
Multiple regression: Cue use in comprehension (β)							
Adverb pretraining-Control (verb use)	–0.01	0.09	54	0.17	0.28	36	–0.70
Verb pretraining-Control (adverb use)	0.76	0.20	54	0.93	0.28	36	–0.55
Group mean correlation: Cue use in comprehension (*r*)							
Adverb pretraining-Control (verb use)	–0.01	0.06	16	0.07	0.29	17	0.41
Verb pretraining-Control (adverb use)	0.46	0.47	17	0.44	0.45	17	0.05
Cue production (*M*)							
Adverb pretraining-Control (verb use)	1.08	1.33	16	2.52	2.67	17	–0.68
Verb pretraining-Control (adverb use)	7.27	3.55	17	5.58	4.51	17	0.42

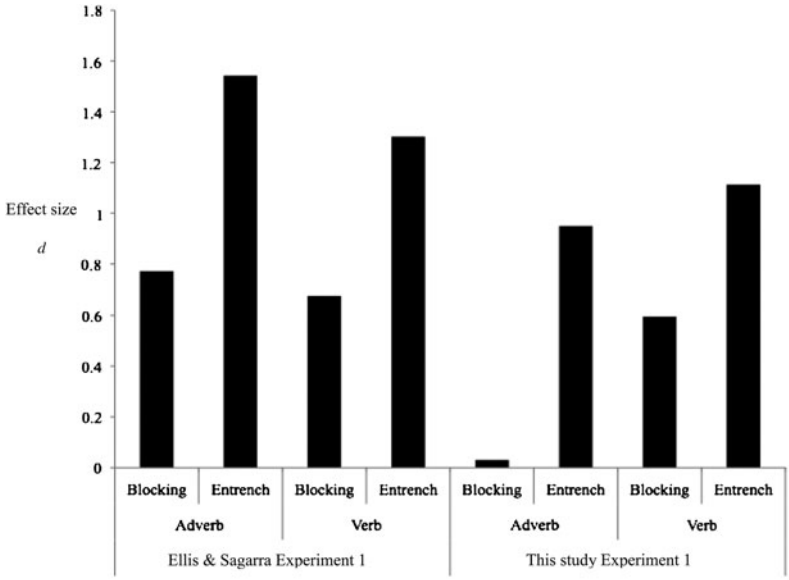


Figure 10. Meta-analysis of effect sizes for entrenchment and blocking as a function of study and cue.

the present Experiment 1 tend not to be blocked by prior experience of the more complicated and less salient morphological cues.

CONCLUSIONS

In the present experiments—in contrast to Ellis and Sagarra (2010)—there was a strong overall bias toward the acquisition of adverbs in control conditions where both cues were always present and where there was no pretraining on any one cue. The stimuli used in Ellis and Sagarra were a meager subset of Latin, where the three adverbs differed from each other in relatively slight ways (*hodie* "today," *heri* "yesterday," *cras* "tomorrow") approximating the similarity of the verbal inflections (*cogito* "I think," *cogitavi* "I thought," *cogitabo* "I will think"). In natural languages, this is not the typical case. Due to its high frequency, verbal morphology is typically of low salience in its surface manifestations compared to lexical cues (*yesterday, today, tomorrow* vs. *I walk**ed***, *I walk*, *I'**ll** walk*); thus, inflections are typically overshadowed by more salient lexical and discourse cues (Bates & Goodman, 1997; Ellis, 2006b, 2006c). In the present experiments, in which the verbal paradigm reflected tense and person in a more naturally complete state, the participants relied much more on the adverbial cues. They were biased

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

toward reliable and salient lexical cues to temporality and this blocked acquisition of less salient and complex morphological cues.

Ellis and Sagarra (2010) acknowledged that their experiments involved language learning in its simplest, most stripped-down form, but they emphasized that they were so designed to allow tight replicable comparisons. The present study demonstrates this replicability: The experiments once again show that the early stages of adult language learning are sensitive to blocking effects from varying sequences of cue exposure and to L1 background (here with a greater range of L1s). Although both of these studies admittedly concern only the first hour of language learning, Sagarra, Ellis, and Gauthier (2011) demonstrated that such effects extend to the sentence processing strategies of third- and eighth-semester English-Spanish FL learners reading sentences in Spanish that contain lexical (adverb) and morphological (verbal inflection) cues to temporal reference.

One key question is why children learning their L1 do not experience this blocking. This issue is considered in some depth in Ellis and Sagarra (2010), through a review of studies suggesting that 2- to 3-year-old children learning English as their L1 focus on morphological means of temporal reference because, at this stage, they do not yet know about pragmatic means of temporal reference (such as recounting events in their order of occurrence), nor do they clearly yet understand the meaning of temporal adverbs. The development of temporal concepts in early childhood is gradual, and there are important changes between 3 and 5 years in the ability to reason about sequences of events in the past and future (see commentary by Ellis, 2008a; McCormack & Hoerl, 2008). Adults, in contrast, have acquired all of these notions by the time they approach their L2. There is scope for detailed comparisons of longitudinal corpora of the development of temporal reference in child L1 and adult L2 learners to more fully investigate these ideas.

The findings of such experiments reinforce the possibility that the limited attainment of adult L2 and FL learning follows general principles of associative learning. Adult FL acquirers are limited in working memory and time on task, and they have attentional biases to language. They know that temporal adverbs are more reliable than nonsalient and ambiguous verbal inflections and that they can usually satisfice (Simon, 1957) and get their message across by lexical means alone—however ungrammatical, the basic variety is communicatively effective.

Together, these experiments demonstrate the following:

1. Short-term positive learned attention effects: FL cues learned early in the instructional sequence become entrenched and are preferentially processed.
2. Short-term negative learned attention effects: FL cues learned early in the instructional sequence block the acquisition of later experienced FL cues to the same interpretation.

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

3. Long-term learned attention effects: The more a learner's L1 makes use of rich verbal morphology, the more they acquire knowledge of FL tense inflections from utterances that contain multiple cues to temporal reference.
4. Cue salience and complexity: The more complicated the FL verbal inflectional paradigm, the more adult learners rely upon salient and simple adverbial cues to time.
5. The magnitude of blocking depends on the competing cues, their relative salience, and the stage of learning. Salient cues resist blocking; less salient and complicated cues are more readily blocked.

(*Received 13 January 2011*)

REFERENCES

Baddeley, A. D. (1976). *The psychology of memory*. New York: Harper and Row.
Bardovi-Harlig, K. (1992). The use of adverbials and natural order in the development of temporal expression. *International Review of Applied Linguistics in Language Teaching*, *30*, 299–320.
Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning, and use*. Oxford: Blackwell.
Barlow, M., & Kemmer, S. (Eds.). (2000). *Usage based models of language*. Stanford, CA: CSLI Publications.
Bates, E., & Goodman, J. C. (1997). On the inseparability of grammar and the lexicon: Evidence from acquisition, aphasia and real-time processing. *Language and Cognitive Processes*, *12*, 507–586.
Cheng, P. W., & Holyoak, K. J. (1995). Adaptive systems as intuitive statisticians: Causality, contingency, and prediction. In J.-A. Meyer & H. Roitblat (Eds.), *Comparative approaches to cognition* (pp. 271–302). Cambridge, MA: MIT Press.
Clahsen, H., & Felser, C. (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, *27*, 3–42.
Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159.
Collins, L., & Ellis, N. C. (Eds.). (2009). Input and second language construction learning: Frequency, form, and function [Special issue]. *Modern Language Journal*, *93*(3).
Comrie, B. (1987). Russian. In B. Comrie (Ed.), *The world's major languages* (pp. 329–347). Oxford: Oxford University Press.
Edwards, A. L. (1984). *An introduction to linear regression and correlation* (2nd ed.). New York: W. H. Freeman.
Ellis, N. C. (1998). Emergentism, connectionism and language learning. *Language Learning*, *48*, 631–664.
Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*, 143–188.
Ellis, N. C. (2003). Constructions, chunking, and connectionism: The emergence of second language structure. In C. Doughty & M. H. Long (Eds.), *Handbook of second language acquisition* (pp. 33–68). Oxford: Blackwell.
Ellis, N. C. (2006a). The associative-cognitive CREED. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 77–96). Mahwah, NJ: Erlbaum.
Ellis, N. C. (2006b). Cognitive perspectives on SLA: The associative-cognitive CREED. *AILA Review*, *19*, 100–121.
Ellis, N. C. (2006c). Selective attention and transfer phenomena in SLA: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, *27*, 1–31.
Ellis, N. C. (2008a). Temporal cognition and temporal language the first and second times around. *Language Learning*, *58*(Suppl. 1), 115–121.

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

Ellis, N. C. (2008b). Usage-based and form-focused language acquisition: The associative learning of constructions, learned-attention, and the limited L2 endstate. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language acquisition* (pp. 372–405). London: Routledge.
Ellis, N. C., & Cadierno, T. (2009). Constructing a second language [Special section]. *Annual Review of Cognitive Linguistics*, *7*, 111–290.
Ellis, N. C., & Sagarra, N. (2010). The bounds of adult language acquisition: Blocking and learned attention. *Studies in Second Language Acquisition*, *32*, 553–580.
Goldschneider, J. M., & DeKeyser, R. (2001). Explaining the "natural order of L2 morpheme acquisition" in English: A meta-analysis of multiple determinants. *Language Learning*, *51*, 1–50.
Green, J. N. (1987). Spanish. In B. Comrie (Ed.), *The world's major languages* (pp. 236–259). Oxford: Oxford University Press.
Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25*, 603–634.
Kamin, L. J. (1969). Predictability, surprise, attention, and conditioning. In B. A. Campbell & R. M. Church (Eds.), *Punishment and aversive behavior* (pp. 276–296). New York: Appleton-Century-Crofts.
Klein, W. (1986). *Second language acquisition*. New York: Cambridge University Press.
Kruschke, J. K. (2006, June). *Learned attention*. Paper presented at the Fifth International Conference on Development and Learning, Bloomington, IN.
Kruschke, J. K., & Blair, N. J. (2000). Blocking and backward blocking involve learned inattention. *Psychonomic Bulletin and Review*, *7*, 636–645.
Lee, J. F. (1999). On levels of processing and levels of comprehension. In J. Gutiérrez-Rexach & F. Martínez-Gil (Eds.), *Advances in Hispanic linguistics* (pp. 42–59). Somerville, MA: Cascadilla Press.
Lee, J. F. (2002). The incidental acquisition of Spanish future morphology through reading in a second language. *Studies in Second Language Acquisition*, *24*, 55–80.
Lee, J. F., Cadierno, T., Glass, W. R., & VanPatten, B. (1997). The effects of lexical and grammatical cues on processing past temporal reference in second language input. *Applied Language Learning*, *8*, 1–23.
Li, C. N., & Thompson, S. A. (1987). Chinese. In B. Comrie (Ed.), *The world's major languages* (pp. 811–833). Oxford: Oxford University Press.
Lin, J.-W. (2006). Time in a language without tense: The case of Chinese. *Journal of Semantics*, *23*, 1–53.
Mackintosh, N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, *82*, 276–298.
MacWhinney, B. (1987). The competition model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249–308). Mahwah, NJ: Erlbaum.
MacWhinney, B. (1997). Second language acquisition and the competition model. In A. M. B. De Groot & J. F. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 113–142). Mahwah, NJ: Erlbaum.
MacWhinney, B. (2001). The competition model: The input, the context, and the brain. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 69–90). New York: Cambridge University Press.
MacWhinney, B., & Bates, E. (1989). *The crosslinguistic study of sentence processing*. New York: Cambridge University Press.
MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, *23*, 127–150.
Mangubhai, F. (1991). The processing behaviors of adult second language learners and their relationship to second language proficiency. *Applied Linguistics*, *12*, 268–297.
Matessa, M., & Anderson, J. R. (2000). Modeling focused learning in role assignment. *Language and Cognitive Processes*, *15*, 263–292.
McCormack, T., & Hoerl, C. (2008). Temporal decentering and the development of temporal concepts. In P. Indefrey & M. Gullberg (Eds.), *Time to speak: Cognitive and neural prerequisites of time in language* (pp. 89–113). Oxford: Blackwell.
McDonald, J. L. (1986). The development of sentence comprehension strategies in English and Dutch. *Journal of Experimental Child Psychology*, *41*, 317–335.

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press

Meisel, J. (1987). Reference to past events and actions in the development of natural second language acquisition. In C. Pfaff (Ed.), *First and second language acquisition* (pp. 206–224). Rowley, MA: Newbury House.
Musumeci, D. (1989). *The ability of second language learners to assign tense at the sentence level: A crosslinguistic study*. Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign.
Noyau, C., Klein, W., & Dietrich, R. (1995). *Acquisition of temporality in a second language*. Amsterdam: Benjamins.
Odlin, T. (1989). *Language transfer*. New York: Cambridge University Press.
R Development Core Team (2008). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.
Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current theory and research* (pp. 64–99). New York: Appleton-Century-Crofts.
Robinson, P., & Ellis, N. C. (Eds.). (2008). *A handbook of cognitive linguistics and second language acquisition*. London: Routledge.
Sagarra, N. (2001, October). *Attention allocation to morphological cues during L2 sentence processing: Evidence from eye-movement*. Paper presented at the Fourth Conference on the Acquisition of Spanish and Portuguese as First and Second Languages, Urbana-Champaign, IL.
Sagarra, N., Ellis, N., & Gauthier, J. (2011, June). *Processing verbal morphological agreement in L1 and L2: Language experience, working memory and linguistic effects.* Paper presented at the Seventh International Morphological Processing Conference, Donostia, Spain.
Schmidt, R. (1984). The strengths and limitations of acquisition: A case study of an untutored language learner. *Language, Learning, and Communication*, *3*, 1–16.
Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). New York: Cambridge University Press.
Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide*. Pittsburgh, PA: Psychology Software Tools Inc.
Shanks, D. R. (1995). *The psychology of associative learning*. New York: Cambridge University Press.
Simon, H. A. (1957). *Models of man: Social and rational*. New York: Wiley.
Terrell, T. (1991). The role of grammar instruction in a communicative approach. *Modern Language Journal*, *75*, 52–63.
Thalheimer, W., & Cook, S. (2002). *How to calculate effect sizes from published research: A simplified methodology*. Retrieved June 29, 2010, from http://work-learning.com/effect_sizes.htm
VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Westport, CT: Ablex.
VanPatten, B. (2006). Input processing. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 115–136). Mahwah, NJ: Erlbaum.
Wills, A. J. (2005). *New directions in human associative learning*. Mahwah, NJ: Erlbaum.
Xiao, R., & McEnery, T. (2004). *Aspect in Mandarin Chinese*. Amsterdam: Benjamins.
Zobl, H., & Liceras, J. M. (1994). Functional categories and acquisition orders. *Language Learning*, *44*, 159–180.

https://doi.org/10.1017/S0272263111000325 Published online by Cambridge University Press