*Applied Psycholinguistics* (2020), **41**, 627–656
doi:10.1017/S0142716420000119

CAMBRIDGE
UNIVERSITY PRESS

ORIGINAL ARTICLE

# Morphosyntactic adaptation in adult L2 processing: Exposure and the processing of case and tense violations

Holger Hopp

Technische Universität Braunschweig
Corresponding author: Email: h.hopp@tu-braunschweig.de

(Received 12 June 2019; revised 08 January 2020; accepted 29 January 2020)

**Abstract**
This paper investigates morphosyntactic adaptation in second language (L2) sentence processing. In a pre-/posttest control group design, two experiments with intermediate to advanced German–English learners examine whether massed exposure to informative input leads to adaptation in L2 processing in that L2 readers come to integrate inflection in real-time comprehension. Experiment 1 on case marking shows that input causing prediction error and flagging the target parse leads to nativelike integration of case in the reanalysis of garden-path sentences. Experiment 2 shows partially nativelike processing of adverbial–verb tense mismatches after exposure to target input. Adaptation was selective to the experimental versus the control group in processing, yet it did not generalize to offline, explicit performance. We conclude that morphosyntactic adaptation constitutes an implicit learning mechanism in L2 processing, and we discuss its implications for models of L2 processing and acquisition.

**Keywords:** adaptation; inflectional variability; morphosyntax; prediction; second language processing

Mastering inflectional morphology proves to be one of the most challenging tasks for adult second language (L2) learners, and even very advanced learners often produce target and nontarget forms variably. Research on L2 sentence processing has been investigating such inflectional variability in real-time comprehension. Many studies on the L2 processing of subject–verb agreement, case marking, or gender agreement report that adult L2 learners also have difficulties integrating inflection incrementally in parsing. Such difficulties have been interpreted as evidence that L2 learners underuse morphosyntactic cues (e.g., Cunnings, 2017; MacWhinney, 1997) or that they are unable to recruit grammatical representations in real-time sentence processing (e.g., Clahsen & Felser, 2006; Jiang, 2007). Most studies consider the processing of inflection at one point in time and then conclude that inflectional variability is a general property of adult L2 processing. However, developmental studies show that difficulties with inflection reduce as learners become more proficient (e.g., Hopp, 2010; Hoshino et al., 2010), receive instruction (e.g., Morgan-Short, Sanz, Steinhauer, & Ullman, 2010) or learn incidentally from the input (e.g., Hopp, 2016). These findings illustrate that the L2 processing of inflection is not static over time and may change in the long run.

Recent research on native-language (L1) processing suggests that processing may also change in the short run. Native speakers quickly adapt to changes in the surrounding language input, for example, differences in accents or vocabulary use across speakers as well as shifting uses of syntactic structures (for review, Kaan & Chun, 2018b). Evidence of syntactic adaptation comes from priming studies as well as reading studies in which infrequent or dispreferred structures accumulate in the course of an experiment. For instance, English natives show decreasing processing difficulty with reduced relative clauses (e.g., Fine, Jaeger, Farmer, & Qian, 2013; though see Harrington-Stack, James, & Watson, 2018) or dispreferred coordination structures (e.g., Kaan, Futch, Fuertes, Mujcinovic, & de la Fuente, 2019) after repeated exposure to these structures in the input. Syntactic adaptation has been taken to reflect error-based implicit learning (e.g., Dell & Chang, 2014) or probabilistic belief updating (e.g., Jaeger & Snider, 2013). To explain adaptation, these approaches assume that speakers make (implicit) predictions about the upcoming input and use prediction error as feedback to restructure their linguistic knowledge and expectations to match the changing statistics in the input so as to minimize future prediction error. In different frameworks, the interplay of input, prediction error, and adaptation has been postulated as a powerful implicit learning mechanism in L1 acquisition and L1 adult processing (e.g., Christiansen & Chater, 2016; Dell & Chang, 2014; Phillips & Ehrenhofer, 2015).

The evidence as to whether (adult) L2 learners also demonstrate syntactic adaptation is mixed. On the one hand, L2 priming studies report that L2 learners show priming effects (e.g., Jackson & Ruf, 2016; Kaan & Chun, 2018a) and that the strength of priming may be comparable in L1 and L2 (e.g., Hartsuiker, Beerts, Loncke, Desmet, & Bernolet, 2016). On the other hand, reading time studies on syntactic adaptation report that adult L2 learners may not adapt to the input like native speakers or that they may need more input than native speakers for adaptation (Kaan et al., 2019; yet see Arai, 2016).

The present study uses the syntactic adaptation paradigm to study inflectional variability in L2 processing. We systematically build on previous studies showing that learners lack sensitivity to morphosyntactic detail in L2 processing. For one thing, the present study thus aims to replicate previous studies. For another, it serves to probe adaptation across two different phenomena and across different methods. We test whether targeted processing experience that aims to highlight prediction error can enhance L2 speakers' sensitivity to inflection, specifically case marking and grammatical tense marking. In the current paper, we replicate two previous studies that have found that L2 learners are not sensitive to case marking (Experiment 1; Hopp, 2014) or tense mismatches (Experiment 2; Roberts & Liszka, 2013) in processing. In a control-group design, the experimental group reads sentences in which the target structure accrues in informative contexts, while a control group reads comparable sentences that do not contain morphosyntactic information regarding the target structure. In a posttest, we test whether learners in the experimental group have adapted to the changes in the input.

The paper is structured as follows. We review the role of prediction error for the implicit learning of inflection in sentence processing by native and nonnative speakers. Then we report Experiment 1 on case marking, and describe Experiment 2 on tense mismatches. We then discuss the findings in the context of current approaches to L2 sentence processing.

## Prediction error and the implicit learning of morphosyntax

Behavioral eye-tracking studies and electrophysiological studies report that native speakers predict at many levels, including inflection signaling agreement (for review, see Kuperberg & Jaeger, 2016; Pickering & Gambi, 2018). For instance, native speakers use case marking (Kamide et al., 2003; Hopp, 2015) and gender marking (e.g., Lew-Williams & Fernald, 2010; Wicha, Moreno, & Kutas, 2004) to anticipate upcoming referents in the input.

On top of prediction increasing the efficiency of conversation and turn taking through speaker–hearer alignment (e.g., Pickering & Garrod, 2013), prediction serves as a powerful learning mechanism in that it enables the learner to generate his or her own feedback, namely, by comparing the unfolding input to the internally generated predictions about the input (e.g., Phillips & Ehrenhofer, 2015, though see Huettig & Mani, 2016). If the two do not match, the learner experiences a prediction error, and she can use the information in the input signaling prediction error to adjust or update her expectations to the input properties.

According to formal accounts of language acquisition, prediction allows the learner to actively explore hypotheses about the target language, and she can use prediction error to revise her hypotheses in order to gradually converge on the target grammar (e.g., Phillips & Ehrenhofer, 2015). In connectionist learning models, prediction error creates feedback loops for the network to approximate the statistical distributions of the input. For instance, in the P-chain model (Dell & Chang, 2014), the learner employs covert production routines to predict the unfolding input, and the learner can use prediction error to adjust the weightings of connections in the network to reduce future prediction error. According to functionalist accounts, language users keep track of the input statistics and form expectations (beliefs) about the input distributions. When the input does not match these expectations, they update their beliefs to optimize communicative efficiency (e.g., Kleinschmidt & Jaeger, 2015). Despite their different theoretical underpinnings, all of these approaches cohere in assigning prediction error a central role in learning (Rabagliati, Gambi, & Pickering, 2015).

In L1 acquisition, morphosyntactic prediction develops early (e.g., Lew-Williams & Fernald, 2007; Lukyanenko & Fisher, 2016; van Heugten & Shi, 2009) and in tandem with production skills (Mani & Huettig, 2012). Further, children use grammatical structure rather than associations for prediction (Gambi, Pickering, & Rabigliati, 2016), and they have been found to adapt their morphosyntactic predictions following prediction error engendered by unexpected input (e.g., Havron, de Carvalho, Fiévet, & Christophe, 2019).

In adult L1 processing, language users equally use prediction error to adapt to changes in the statistics in the input. For instance, German natives adapt to changing statistics of inflectional or prosodic cues that signal word order (Henry, Hopp, & Jackson, 2017). In addition, they adapt to how different speakers use inflection to signal word order (Kroczek & Gunter, 2017; see also Kamide, 2012). These speaker-specific predictions persevered over a period of 9 months, indicating that morphosyntactic adaptation can result in long-term implicit learning (see also Kaschak, Kutta, & Coyle, 2014).

In contrast, adult L2 learners engage much less in morphosyntactic prediction during processing (Grüter, Rohde, & Schafer, 2017; Kaan, 2014). Although they readily use semantic and lexical cues in predictive processing (e.g., Ito, Pickering, & Corley, 2017), L2 learners show lower or no predictive processing of gender marking (e.g., Hopp, 2013; Lew-Williams & Fernald, 2010; yet see Dussias, Valdés Kroff, Guezzardo Tamargo, & Gerfen, 2013) or case marking (e.g., Hopp, 2015; Mitsugi & MacWhinney, 2016), especially if the L1 lacks gender or case. More importantly, even when they encounter a prediction error, L2 speakers do not revise their predictions and continue to commit to erroneous predictions (Hopp, 2015). These findings suggest that, unlike child L1 learners and adult L1 speakers, L2 learners do not easily integrate information that could be used to revise predictions. However, it is an open question whether L2 learners are largely insensitive to prediction error or whether they require more input to adjust their predictions, to learn from prediction error, and to adapt to the target use of morphosyntax.

A study by Hopp (2016) addressed the role of systematic input for strengthening L2 predictive processing in a training study. In a pre-/posttest design, 34 L1 English intermediate learners of L2 German completed a visual-world eye-tracking study testing if L2 learners could use a gender-marked article to anticipate an upcoming noun (e.g., $der_{MASC}$ → $Tisch_{MASC}$ – *the table*). Even when controlling for knowledge of the grammatical gender of the experimental items, the pretest found that the L2 speakers did not show predictive use of gender. A week later, they took part in a 15-min training that repeatedly exposed them to the article-noun sequences (e.g., *der Tisch*) used in the experiment, and they practiced them to criterion. Subsequently, they demonstrated target predictive use of grammatical gender in the posttest eye-tracking study. Hence, massed input and practice of the grammatical gender marking led participants to learn these relations and adapt their processing. While this study underscores that adult L2 learners can come to adapt to predictive processing, this may partially be the result of the item-specific explicit training on gender, so that it remains an open question whether target processing would generalize beyond the particular lexical items trained in the experiment.

In this paper, we report two experiments that test morphosyntactic adaptation in L2 learners. We ask the general research question of whether systematic processing experience enhances L2 speakers' sensitivity to inflection. Experiment 1 addresses the importance of massed exposure for implicit learning in the L2 processing of case marking. It tests how input designed to induce prediction error and to provide a subsequent cue for the reanalysis of a garden-path sentence can yield targetlike L2 processing of case. In short, Experiment 1 tests if input intended to engender and resolve prediction error changes predictions in L2 processing. However, previous studies report that L2 speakers generally engage *less* in predictive processing than monolinguals, so that the target processing of inflection may predominantly require the generation of predictive processing rather than the augmentation of prediction error in L2 processing. In consequence, Experiment 2 assesses the degree to which massed input can bring about prediction in the L2 processing of (mis-) matches between temporal adverbials and grammatical tense marking on verbs. Together, these experiments yield evidence on how adaptation scopes over different morphosyntactic phenomena and learner groups.

## Experiment 1: Prediction error and the use of case marking for reanalysis

Experiment 1 builds on the study by Hopp (2014) that tested the integration of different types of information in the L2 processing of garden-path sentences. The sentences in (1) give rise to a temporary ambiguity of the postverbal noun phrase *the boy/he/the piano* between being the object of the verb *play* or the subject of the main clause. Readers initially predict the postverbal noun to be the object, until the verb *made* forces them to reanalyze the noun as the subject of the main clause.

1. a. When the girl was praying, the boy made some funny noises. (*Control*)
   b. When the girl was praying the boy made some funny noises. (*Intransitive*)
   c. When the girl was playing he made some funny noises. (*Case*)
   d. When the girl was playing the boy made some funny noises. (*Implausible*)
   e. When the girl was playing the piano made some funny noises. (*Plausible*)

In Hopp (2014), native speakers of English showed garden-path effects by virtue of heightened reanalysis effort on the main clause verb *made* for implausible (1d) and plausible (1e) noun phrases; yet, they did not differ in reading times between the intransitive (1b) and the pronoun (1c) condition compared to the control condition (1a). These findings resonate with findings from other studies showing that native speakers integrate subcategorization and case information in parsing (e.g., Staub, 2007; Traxler & Pickering, 1996). Adult L1 German L2 learners of English demonstrated analogous processing patterns, that is, garden-path effects for plausible and implausible noun phrases (1d and 1e) and no difficulty for intransitive verbs (1b); yet, the L2 group did not appear to integrate case information, since reading times for (1c) did not differ from (1d) but were significantly longer than in the control condition (1a). These findings align with other studies that report that intermediate to advanced L2 learners do not make incremental use of case marking for word order revisions (e.g., Hopp, 2006; Jackson & Dussias, 2009; see also Hopp, 2015; Mitsugi & MacWhinney, 2016).

The study by Hopp (2014) lends itself to the investigation of implicit learning through prediction error since the optionally transitive verb *playing* gives rise to the expectation that a postverbal noun phrase is the object of the verb, and readers will integrate it as such in the unfolding parse. When they encounter the main clause verb, they experience a prediction error in that they are forced to reanalyze the noun phrase as a subject. Experiment 1 examines the extent to which additional massed input that consistently induces such prediction error and that points readers to the correct analysis engenders adaptation to target processing of case marking. We compare changes in processing patterns of the experimental group to those of a control group that receives lexically comparable input that does, however, not induce prediction error or flag reanalysis.

**Table 1.** Experiment 1: Participant information, means (standard deviation)—all participants ($n = 82$)

| | Experimental group | Control group |
|---|---|---|
| *N* | 38 [42] | 38 [40] |
| LexTale score | 78.6 (12.3) | 74.7 (12.1) |
| Length of exposure (yr) | 13.0 (2.8) | 12.6 (2.3) |
| Age of onset (yr) | 7.8 (1.5) | 8.0 (1.7) |

We make the following hypotheses and predictions. If learners adapt as a consequence of additional exposure to garden-path sentences, we predict that the processing patterns will change for the experimental group, yet not for the control group. Further, if adaptation is specific to the exposure highlighting the target use of case marking, we predict that learners will only demonstrate processing changes for the case condition (1c), yet not for other garden-path sentences involving prediction error, that is, plausible and implausible sentences (1d and 1e).

## *Participants*

Eighty-two L1 German intermediate to advanced late L2 learners (61 female, age 18–32 years, mean age: 20.8 years) took part in the study. All participants had started learning English later than age 5, and they were students of English at a German university at the time of testing and they had normal or corrected-to-normal vision. They participated voluntarily in exchange for course credit or compensation of 7.50 Euros. Group assignment to the experimental and control group was random. Additional participant information is given in Table 1.

The participants took the LexTALE task (Lemhöfer & Broersma, 2012), and the results placed them into the upper-intermediate to advanced range. One-way analyses of variance (ANOVAs) for length of exposure, age of onset and LexTALE scores with the between-subject factor group did not show any significant differences in any measure (all $ps > .17$).

## *Methods and materials*

### *Pretest and posttest*

The materials were adopted from Hopp (2014) that based the selection of verbs on ratings in Traxler (2002). To reduce the number of conditions, the intransitive condition (1b) was dropped, and minor changes were made to the stimuli, for example, the verb in the control condition was the same as in all other conditions. In addition to the 28 quadruplets adapted from Hopp (2014), 28 comparable quadruplets of sentences were designed using the same optionally transitive verbs but otherwise different lexical items (see online-only Supplemental materials). In all sentences, an optionally transitive verb was followed by a full noun phrase introducing a new referent or a nonanaphoric nominative pronoun. In addition, 56 filler sentences were constructed so that all consisted of sentences involving clausal coordination with two subject-initial clauses with transitive verbs each. In consequence, participants encountered 56 experimental sentences containing one (optionally) intransitive verb and one transitive verb as well as 56 filler sentences containing two transitive verbs each, that is, 56 intransitive and 168 transitive verbs in total. Half of the experimental sentences and half of the fillers were followed by comprehension questions. The comprehension questions for the experimental items targeted the subject of the main clause predicate (e.g., *Who made funny noises?*) as the answer, and the participants were presented the subject of the adjunct clause (e.g., *the girl*) and the subject of the main clause (e.g., *the boy*) as possible answers. This way it was possible to test whether participants ended up construing the noun phrase (NP) as the subject of the main clause.

The 56 experimental items were split into two sets for which four lists were created according to a Latin square design. The sets were assigned to the pretest and the posttest in alternation, and each participant received a different list in the pretest and the posttest. In consequence, each participant encountered the same verb used as an intransitive verb twice, yet in a different condition.

### Exposure phase

For the exposure phase, 28 pairs of sentences were constructed with the same verbs used in the experimental sentences (2).

2. a. The boy played and he pleased the parents with the music. *(Experimental group)*
   b. The boy played the music and it pleased the parents. *(Control group)*

The experimental group read sentences as in (2a), in which the optionally transitive verb of the first clause (i) was used intransitively and (ii) the subject pronoun in the second clause was coreferential with the subject of the first clause and (iii) bore unambiguous nominative case marking signaling it is the subject of the second clause. For the sentence in (2a), readers predict the verb *played* to be followed by an object. Upon encountering *he* following the coordinator *and*, they experience a prediction error and need to adopt an intransitive interpretation of the verb, since *he* unambiguously constitutes a subject. The control group read sentences as in (2b), in which the verb of the first clause (i) was used transitively and (ii) the subject pronoun in the second clause was ambiguous in case marking and (iii) was coreferential with either the subject or the object of the first clause.[1] The control group thus did not experience a prediction error, as the verb *played* was followed by an object.

## Procedure

After signing the consent form, the participants first completed the pretest, then took the LexTALE task, filled in the language background questionnaire and did other tasks unrelated to this study. They then read the sentences in the exposure phase and subsequently took the posttest. In the pre- and posttest as well as the exposure phase, all sentences were presented in pseudorandom order, with a new pseudorandomization for each participant. Instructions and three practice sentences including comprehension questions preceded the main experiment. Reading time and response data were collected using an SMI RED High-Speed eye-tracker with a spatial resolution below 0.4 degrees. Tracking speed was 500 Hz, and participants rested their head on a chin-rest to minimize head movements. The sentences were presented in 20-point Arial in white on black on a 22-inch TFT screen. Participants sat in front of the screen at a distance of 70 cm. Participants were instructed to read at their normal reading speeds for comprehension. Before the first item in the main study, participants were calibrated with a 9-point calibration and a subsequent 4-point validation. The calibration procedure was repeated if visual acuity was below 0.5 degrees. Participants were recalibrated at various points in the experiment if necessary. In all, the participants took between 12 and 20 min to complete the pretest and between 22 and 28 min to complete the combined exposure phase and posttest.

Following Hopp (2014), we defined four regions of interest in the experimental sentences: the verb of the adjunct clause (*playing*), the postverbal noun (*the piano*), the verb of the main clause (*made*), and the postverbal region, that is, the rest of the sentence (*some funny noises*). For each region of interest, we computed first-pass reading time, second-pass reading time, total reading time, and number of regressions, as in Hopp (2014).

### *Analysis and results*

Six participants were excluded from analysis, 4 from the experimental group and 2 from the control group, either because they did not have German as the L1 ($n = 4$), because of large amounts of missing data ($n = 1$), or because of poor calibration accuracy ($n = 1$). For all of the remaining 76 participants, reading times for a particular region were excluded if fixations were shorter than 80 ms or longer than 2000 ms (Rayner, 1998). In all, less than 8.2% of all data were thus excluded. The analysis of the comprehension questions showed that accuracy was higher than 85% in each of the conditions in each group and each test, which indicates that participants read the sentences attentively and that they successfully reanalyzed the garden paths.

Table 2 lists the reading times for each condition and for each region of interest by test and group. For the main clause verb and the postverbal regions, we report planned comparisons between the conditions based on Hopp (2014) for second-pass reading times, total reading times and the number of regressions, that is, measures associated with reanalysis processes (e.g., Staub & Rayner, 2007). First, to test whether readers get garden-pathed at all, we compare (1a) and (1d), which we refer to as the *control comparison*. Readers should show longer reading times in (1d) than in (1a). Second, to establish if readers are sensitive to plausibility differences, we compare (1d) and (1e), which we term the *plausibility comparison*. If readers use plausibility, then the commitment to an object interpretation of the postverbal noun phrase should be lower in (1d), which, in turn, facilitates reanalysis to a subject interpretation compared to a more plausible object noun phrase in (1e). Third, we compare (1a) and (1c), that is, the *case comparison*. If case marking is used incrementally, the postverbal pronoun should not be interpreted as the object and therefore not necessitate reanalysis once the reader encounters the main clause verb.

**Table 2.** Experiment 1: Reading times (in ms) and number of regressions by region and by group ($n = 38$); standard deviations in parentheses

| | Group: | | Experimental group | | | | Control group | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Region | Condition | Test | First Pass | Second Pass | Total reading time | No. of regressions | First Pass | Second Pass | Total reading time | No. of regressions |
| Adjunct clause verb | Control (1a) | Pretest | 359 (112) | 321 (205) | 680 (251) | 117 | 293 (86) | 218 (149) | 510 (173) | 118 |
| | Control (1a) | Posttest | 331 (97) | 175 (153) | 506 (199) | 82 | 289 (112) | 154 (128) | 442 (177) | 69 |
| | Case (1c) | Pretest | 356 (79) | 320 (203) | 676 (241) | 88 | 285 (85) | 246 (149) | 531 (186) | 90 |
| | Case (1c) | Posttest | 331 (77) | 254 (223) | 585 (257) | 76 | 313 (108) | 203 (134) | 516 (177) | 64 |
| | Implausible (1d) | Pretest | 363 (112) | 359 (248) | 722 (301) | 105 | 308 (91) | 332 (197) | 640 (235) | 98 |
| | Implausible (1d) | Posttest | 339 (92) | 261 (248) | 599 (316) | 83 | 306 (118) | 192 (132) | 498 (193) | 59 |
| | Plausible (1e) | Pretest | 348 (112) | 394 (273) | 743 (317) | 94 | 297 (87) | 361 (230) | 658 (256) | 98 |
| | Plausible (1e) | Posttest | 352 (117) | 279 (284) | 630 (325) | 76 | 295 (95) | 259 (201) | 554 (234) | 71 |
| NP | Control (1a) | Pretest | 392 (119) | 273 (186) | 665 (244) | 68 | 402 (114) | 257 (154) | 659 (181) | 98 |
| | Control (1a) | Posttest | 355 (87) | 178 (173) | 533 (205) | 49 | 380 (98) | 195 (123) | 576 (168) | 72 |
| | Case (1c) | Pretest | 286 (83) | 157 (121) | 443 (171) | 50 | 248 (77) | 154 (139) | 402 (152) | 57 |
| | Case (1c) | Posttest | 257 (62) | 119 (133) | 376 (153) | 50 | 276 (75) | 126 (106) | 402 (141) | 54 |
| | Implausible (1d) | Pretest | 422 (136) | 382 (287) | 804 (346) | 98 | 429 (121) | 390 (221) | 819 (245) | 133 |
| | Implausible (1d) | Posttest | 444 (126) | 252 (223) | 695 (280) | 68 | 416 (92) | 257 (141) | 673 (167) | 83 |
| | Plausible (1e) | Pretest | 402 (127) | 461 (268) | 862 (343) | 116 | 393 (98) | 412 (244) | 805 (248) | 129 |
| | Plausible (1e) | Posttest | 398 (140) | 307 (252) | 705 (295) | 74 | 380 (96) | 309 (157) | 690 (207) | 100 |

*(Continued)*

**Table 2.** *(Continued)*

| | Group: | | Experimental group | | | | Control group | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Region | Condition | Test | First Pass | Second Pass | Total reading time | No. of regressions | First Pass | Second Pass | Total reading time | No. of regressions |
| Main verb | Control (1a) | Pretest | 318 (99) | 167 (151) | 486 (191) | 64 | 332 (92) | 158 (112) | 490 (136) | 66 |
| | Control (1a) | Posttest | 299 (87) | 112 (96) | 412 (145) | 54 | 319 (73) | 144 (113) | 463 (129) | 58 |
| | Case (1c) | Pretest | 350 (115) | 231 (153) | 581 (181) | 81 | 330 (118) | 227 (165) | 558 (179) | 109 |
| | Case (1c) | Posttest | 294 (93) | 157 (131) | 452 (178) | 86 | 380 (104) | 188 (153) | 568 (202) | 89 |
| | Implausible (1d) | Pretest | 350 (110) | 248 (206) | 598 (250) | 101 | 354 (100) | 255 (151) | 609 (171) | 143 |
| | Implausible (1d) | Posttest | 321 (83) | 184 (173) | 505 (212) | 90 | 347 (91) | 172 (119) | 519 (157) | 87 |
| | Plausible (1e) | Pretest | 340 (98) | 369 (237) | 709 (279) | 134 | 361 (93) | 336 (225) | 698 (232) | 131 |
| | Plausible (1e) | Posttest | 334 (94) | 229 (199) | 563 (229) | 107 | 360 (111) | 214 (169) | 574 (214) | 109 |
| Post-verbal | Control (1a) | Pretest | 585 (219) | 166 (150) | 752 (245) | 294 | 559 (216) | 160 (142) | 719 (256) | 267 |
| | Control (1a) | Posttest | 526 (174) | 103 (104) | 629 (232) | 224 | 510 (176) | 103 (92) | 613 (211) | 234 |
| | Case (1c) | Pretest | 598 (181) | 246 (205) | 845 (281) | 305 | 519 (192) | 216 (149) | 735 (241) | 288 |
| | Case (1c) | Posttest | 549 (178) | 123 (126) | 672 (227) | 236 | 532 (183) | 142 (116) | 674 (207) | 252 |
| | Implausible (1d) | Pretest | 530 (165) | 243 (213) | 773 (272) | 309 | 521 (193) | 211 (167) | 731 (246) | 283 |
| | Implausible (1d) | Posttest | 517 (155) | 132 (147) | 649 (243) | 251 | 518 (168) | 124 (130) | 642 (214) | 216 |
| | Plausible (1e) | Pretest | 529 (161) | 346 (223) | 875 (298) | 347 | 462 (163) | 330 (187) | 792 (260) | 342 |
| | Plausible (1e) | Posttest | 535 (165) | 197 (173) | 732 (271) | 265 | 467 (157) | 191 (167) | 657 (220) | 254 |

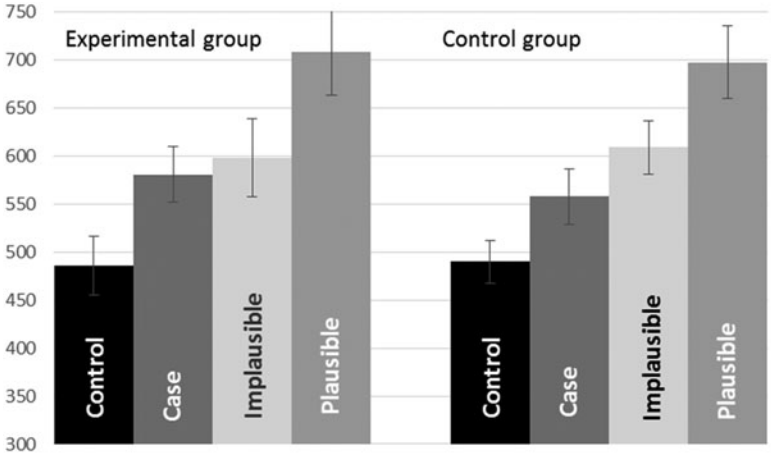

**Figure 1.** Experiment 1: Pretest. Mean total reading times (in ms) on the main verb region by group ($n = 38$ each). Error bars show standard error.

Analyses of all reading time measures were conducted using mixed-effect regression models with the lme4 and lmerTest packages in *R* version 3.5.1 (R Development Core Team, 2018). Effect sizes for comparisons with one fixed effect and random factors were computed following Westfall, Kenny, and Judd (2014). To afford comparability with the study by Hopp (2014), we analyed raw reading times.[2] We entered Group (experimental vs. control), Condition (1a, c, d and e, for each comparison), Test (pretest vs. posttest) as fixed factors, including their interactions. We also used Proficiency (LexTALE score) and Trial number as continuous and centered fixed effects including their interactions with Condition and Test. The random effects structure included random intercepts for participants and items, as well as all random slopes justified by the design. When this maximal model did not converge, we first removed the by-item, and then the by-participant random correlation parameters (see Barr, Levy, Scheepers, & Tily, 2013, p. 276). All converging models minimally contained random intercepts for participants and items.

As in Hopp (2014), Figures 1 and 2 plot the total reading times on the main verb region by group for the pretest and the posttest, respectively. The results of the pretest in each group replicate the findings for the L2 learners in Hopp (2014), in that each group showed main effects of Condition for the *control comparison*, the *plausibility comparison*, as well as the *case comparison* (Table 3), that is, each group showed garden-path effects in all conditions.

To test adaptation, we are interested in seeing whether there is an interaction of Condition and Test, which would indicate that reading time patterns change from pretest to posttest for each comparison. We first turn to the *control* and the *plausibility comparisons*. Models including Group revealed no significant interactions of Group, Condition, and Test for the control or the plausibility comparisons. Follow-up analyses by group did not show any significant interactions of Condition and Test for any measures in any group. In addition, the reading time patterns remained comparable in

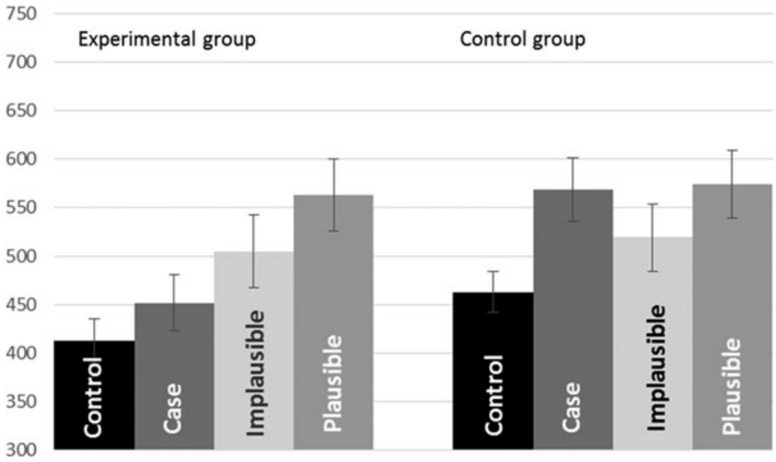

**Figure 2.** Experiment 1: Posttest. Mean total reading times (in ms) on the main verb region by group ($n = 38$ each). Error bars show standard error.

pretest and posttest (see Table 3). This shows that the exposure phase did not cancel the garden-path effects and the plausibility effects in either group.

For the *case comparison*, we did not find any significant interactions in first-pass or second-pass reading times or the number of regressions; yet, there was a marginally significant interaction of Group, Condition, and Test ($\beta = 113.68$, $SE = 61.01$, $t = 1.863$, $p = .062$) for total reading times in the postverbal region. Hence, we analyzed the data of the experimental and the control groups separately in order to check if both groups demonstrated the same pattern of effects.

For second-pass reading times in the postverbal region, the interaction between Condition and Test became significant ($\beta = 16.91$, $SE = 8.61$, $t = 1.965$, $p = .049$). When looking at the effect patterns across pretest and posttest in the experimental group, the significant main effects of Condition in the pretest disappeared for total reading times on the main verb region as well as for second-pass and total reading times in the postverbal region (Table 3), with effect sizes consistently reducing. For the control group, the models did not yield any effect of Condition (Table 3), and Figures 1 and 2 illustrate that the mean total reading times remain virtually identical from pretest to posttest. In addition, the effect sizes in the control group did not change or they even increased, bearing out that the control group continued to have difficulty with case-marked pronouns.

To substantiate the group differences for case-marked pronouns, we directly compared the total reading times on the main verb region between the groups in the case condition (1c; Figures 1 and 2). A model with the fixed factors Group and Test and the same random effects components as above returned a main effect of Test ($\beta = 121.73$, $SE = 30.09$, $t = 4.046$, $p < .001$) as well as a significant interaction of Test and Group ($\beta = 116.26$, $SE = 42.37$, $t = 2.744$, $p = .008$). This interaction underlines that the change in reading times from pretest to posttest was specific to the experimental group (Table 3). For neither group nor any comparison did L2 proficiency yield any significant interactions. Further, Trial number

**Table 3.** Experiment 1: Model comparisons—main effects of Condition by comparison, region, and group ($n = 38$ each)

| | | Experimental group | | Control group | |
|---|---|---|---|---|---|
| Control comparison (1a vs. 1d) | | Pretest | Posttest | Pretest | Posttest |
| Main verb region | First-pass RT | β: 31.8; *SE*: 17.6; *t*: 1.808+; *d*: 0.15 | β: 18.4; *SE*: 13.8; *t*: 1.335; *d*: 0.11 | β: 26.2; *SE*: 15.1; *t*: 1.738+; *d*: 0.13 | β: 28.6; *SE*: 14.9; *t*: 1.923+; *d*: 0.15 |
| | Second-pass RT | β: 41.6; *SE*: 13.8; *t*: 3.020**; *d*: 0.29 | β: 35.5; *SE*: 10.7; *t*: 3.332***; *d*: 0.27 | β: 98.0; *SE*: 22.8; *t*: 4.297***; *d*: 0.36 | β: 27.6; *SE*: 21.0; *t*: 1.312; *d*: 0.11 |
| | Total RT | β: 57.6; *SE*: 14.9; *t*: 3.878***, *d*: 0.33 | β: 45.3; *SE*: 12.2; *t*: 3.708***, *d*: 0.28 | β: 124.2; *SE*: 26.3; *t*: 4.732***; *d*: 0.38 | β: 57.3; *SE*: 24.4; *t*: 2.346*; *d*: 0.18 |
| Postverbal region | First-pass RT | β: 58.9; *SE*: 26.6; *t*: 2.217*; *d*: 0.17 | β: 3.7; *SE*: 23.9; *t*: 0.157; *d*: 0.01 | β: 41.5; *SE*: 26.0; *t*: 1.597; *d*: 0.12 | β: 8.3; *SE*: 24.2; *t*: 0.341; *d*: 0.03 |
| | Second-pass RT | β: 38.1; *SE*: 12.6; *t*: 3.037**; *d*: 0.24 | β: 15.9; *SE*: 11.6; *t*: 1.366; *d*: 0.12 | β: 50.7; *SE*: 26.5; *t*: 1.917+; *d*: 0.16 | β: 18.8; *SE*: 21.2; *t*: 0.889; *d*: 0.08 |
| | Total RT | β: 8.7; *SE*: 19.3; *t*: 0.451; *d*: 0.04 | β: 15.5; *SE*: 13.7; *t*: 1.130; *d*: 0.08 | β: 9.8; *SE*: 30.7; *t*: 0.320; *d*: 0.03 | β: 29.0; *SE*: 26.3; *t*: 1.104; *d*: 0.08 |
| Plausibility comparison (1d vs. 1e) | | | | | |
| Main verb region | First-pass RT | β: 9.3; *SE*: 16.4; *t*: 0.564; *d*: 0.05 | β: 18.9; *SE*: 15.3; *t*: 1.233; *d*: 0.09 | β: 6.5; *SE*: 18.9; *t*:0.386; *d*: 0.03 | β: 12.1; *SE*: 16.8; *t*: 0.718; *d*: 0.05 |
| | Second-pass RT | β: 50.5; *SE*: 15.0; *t*: 3.366**; *d*: 0.33 | β: 19.1; *SE*: 11.6; *t*: 1.651+; *d*: 0.15 | β: 69.7; *SE*: 26.4; *t*: 2.641**; *d*: 0.22 | β: 39.1; *SE*: 22.2; *t*: 1.760+; *d*: 0.14 |
| | Total RT | β: 53.5; *SE*: 16.2; *t*: 3.302**; *d*: 0.26 | β: 32.2; *SE*: 14.9; *t*: 2.154*; *d*: 0.17 | β: 75.7; *SE*: 29.5; *t*: 2.564*; *d*: 0.22 | β: 51.0; *SE*: 25.6; *t*: 1.996*; *d*: 0.14 |
| Postverbal region | First-pass RT | β: 0.2; *SE*: 26.5; *t*: 0.009; *d*: 0.00 | β: 18.1; *SE*: 24.5; *t*: 0.740; *d*: 0.06 | β: 56.5; *SE*: 26.1; *t*: 2.163*; *d*: 0.17 | β: 56.8; *SE*: 25.5; *t*: 2.225*; *d*: 0.17 |
| | Second-pass RT | β: 54.6; *SE*: 16.3; *t*: 3.344***; *d*: 0.27 | β: 32.5; *SE*: 14.3; *t*: 2.285*; *d*: 0.20 | β: 109.3; *SE*: 29.9; *t*: 3.653***; *d*: 0.30 | β: 69.6; *SE*: 25.5; *t*: 2.726**; *d*: 0.24 |

(*Continued*)

**Table 3.** (*Continued*)

| | | Experimental group | | Control group | |
|---|---|---|---|---|---|
| Control comparison (1a vs. 1d) | | Pretest | Posttest | Pretest | Posttest |
| | Total RT | β: 52.6; *SE*: 18.3;<br>*t*: 2.869**; *d*: 0.21 | β: 42.0; *SE*: 16.1;<br>*t*: 2.613**; *d*: 0.21 | β: 52.3; *SE*: 31.7;<br>*t*: 1.650; *d*: 0.12 | β: 12.0; *SE*: 28.9;<br>*t*: 0.415; *d*: 0.03 |
| Case comparison (1a vs. 1c) | | | | | |
| Main verb region | First-pass RT | β: 26.5; *SE*: 18.0;<br>*t*: 1.467; *d*: 0.12 | β: 18.4; *SE*: 13.8;<br>*t*: 1.335; *d*: 0.11 | β: 2.0; *SE*: 15.9;<br>*t*:0.128; *d*: 0.01 | β: 59.7; *SE*: 18.4;<br>*t*: 3.253**; *d*: 0.27 |
| | Second-pass RT | β: 31.4; *SE*: 13.0;<br>*t*: 2.417*; *d*: 0.23 | β: 22.3; *SE*: 10.1;<br>*t*: 2.203*; *d*: 0.21 | β: 70.0; *SE*: 22.2;<br>*t*: 3.162*; *d*: 0.26 | β: 41.1; *SE*: 21.9;<br>*t*: 1.879+; *d*: 0.16 |
| | Total RT | **β: 45.1; *SE*: 14.2;<br>*t*: 3.167**; *d*: 0.25** | **β: 19.8; *SE*: 13.1;<br>*t*: 1.510; *d*: 0.15** | β: 71.3; *SE*: 25.4;<br>*t*: 2.804**; *d*: 0.22 | β: 99.6; *SE*: 27.5;<br>*t*: 3.618***; *d*: 0.29 |
| Postverbal region | First-pass RT | β: 13.1; *SE*: 28.6;<br>*t*: 0.457; *d*: 0.04 | β: 22.2; *SE*: 23.6;<br>*t*: 0.937; *d*: 0.07 | β: 47.1; *SE*: 26.8;<br>*t*: 1.761+; *d*: .13 | β: 23.1; *SE*: 23.6;<br>*t*: 0.976; *d*: 0.07 |
| | Second-pass RT | **β: 41.1; *SE*: 13.5;<br>*t*: 3.057**; *d*: 0.25** | **β: 7.1; *SE*: 10.6;<br>*t*: 0.671; *d*: 0.06** | β: 60.9; *SE*: 25.7;<br>*t*: 2.396*; *d*: 0.20 | β: 40.3; *SE*: 20.6;<br>*t*: 1.955+; *d*: 0.17 |
| | Total RT | **β: 48.8; *SE*: 17.5;<br>*t*: 2.782**; *d*: 0.21** | **β: 18.0; *SE*: 13.4;<br>*t*: 1.339; *d*: 0.10** | β: 16.1; *SE*: 29.3;<br>*t*: 0.549; *d*: 0.04 | β: 62.6; *SE*: 26.0;<br>*t*: 2.403*; *d*: 0.17 |

*Note*: The values in the cells are coefficient estimate β/standard error *SE*(β) and the associated *t* score. $+p < .10$, $^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

**Table 4.** Experiment 1: Comprehension accuracy (in %) by condition, test, and by group ($n = 38$)

| | | Experimental group | Control group |
|---|---|---|---|
| Control (1a) | Pretest | 95.3 (8.6) | 93.7 (12.4) |
| Control (1a) | Posttest | 95.3 (9.8) | 96.9 (8.6) |
| Case (1c) | Pretest | 87.8 (16.6) | 86.8 (14.2) |
| Case (1c) | Posttest | 92.1 (14.4) | 84.7 (16.4) |
| Implausible (1d) | Pretest | 92.6 (11.8) | 93.4 (9.9) |
| Implausible (1d) | Posttest | 95.3 (8.6) | 96.3 (9.1) |
| Plausible (1e) | Pretest | 91.6 (12.8) | 93.7 (9.4) |
| Plausible (1e) | Posttest | 91.6 (10.0) | 92.1 (11.9) |

did not interact with Condition in the pre- or posttest in any comparison for any group.

Finally, we consider differences in comprehension accuracy, that is, the ability to reanalyze and interpret the sentence correctly. Table 4 lists the mean comprehension accuracy by group, condition, and test.

We carried out logistic mixed-effects regressions with Condition, Test, and Group as fixed factors as well as random intercepts for participants and items, and random slopes for Test for both participants and item. For the *control* or the *plausibility comparison*, the models yielded no main effects of Test or any interactions. For the *case comparison*, though, there was a marginally significant three-way interaction between Condition, Test, and Group ($\beta = 0.19$, $SE = 0.10$, $t = 1.820$, $p = .069$). As seen in Table 4, the experimental group improved in comprehension accuracy from pretest to posttest, while the control group did not. In sum, then, the differences between groups in reading times for the case comparison were mirrored in comprehension accuracy, that is, the ability to reanalyze the sentences to the correct interpretation.

### *Experiment 1: Discussion*

Experiment 1 tested whether exposure to sentences that induce prediction error and flag nominative case-marked pronouns as subjects leads L2 learners to integrate case marking in their processing of garden-path sentences.

In the pretest, the findings replicate the results in Hopp (2014) in that L2 learners experience reanalysis effort in sentences with unambiguously nominative pronouns in the postverbal position. Hence, for L2 learners, nominative case marking did not attenuate the integration of a postverbal pronoun as an object, and it did not aid in the reanalysis process toward the pronoun being the subject of the main clause. Reading times on the main verb region for the pronoun condition (1c) were comparable to those for implausible objects (1d), which indicates that grammatical form did not alleviate the initial object misanalysis. In addition, the present study also replicates that L2 learners made robust use of plausibility, with nouns that are plausible objects to the adjunct clause verb incurring greater processing effort than nouns that constitute implausible objects (1d vs. 1e). In terms of adaptation, this means that experiencing frequent prediction errors of optionally transitive verbs being used intransitively in the experimental items did not in and of itself lead learners to adapt to this structure (see also Andrews, Dillon, & Staub, 2019, for native speakers).

Instead, adaptation was specific to the experimental group that read sentences in the exposure phase in which they would encounter more sentences inducing prediction errors due to the intransitive use of optionally transitive verbs. Yet the systematic induction of prediction error in the exposure phase in the experimental group did not lead learners to adapt to an intransitive reading of the adjunct clause verbs in the posttest in general. If so, the group would have shown a reduction of the garden-path effects in the control comparison and the plausibility comparison as well, since the first verb is used intransitively in all of these conditions. Rather, the group specifically adapted to the nominative-marked pronoun signaling a subject of a new clause. The specificity of the adaptation effects in Experiment 1 suggests that merely encountering a prediction error is not sufficient for adaptation in L2 learners (see Kaan, 2014). In addition, learners need to receive information as to what the intended reading is; in Experiment 1, such information was provided by the nominative-marked pronoun that flags the pronoun as the subject of the main clause. The findings from Experiment 1 suggest that L2 learners adapt if prediction error is accompanied by information how the reader can construct the correct parse (see Hopp, 2006). In this case, readers come to show less garden-pathing and greater comprehension accuracy in the posttest.

The findings from Experiment 1 thus confirm both hypotheses in that adaptation was limited to (i) the experimental group and (ii) to sentences involving case-marked pronouns. Accordingly, massed input can lead L2 learners to integrate case marking as signaling prediction error. As reviewed above, however, adult L2 learners routinely do not make morphosyntactic predictions in the first place, so that a crucial next question is whether massed and targeted input can lead to adult L2 learners developing target predictions. To this end, Experiment 2 investigates a different phenomenon to test if L2 learners can come to make predictions involving inflectional morphology, namely, tense marking.

## EXPERIMENT 2

Experiment 2 builds on the study by Roberts and Liszka (2013) that investigated sensitivity to tense mismatches in the L2 processing of English. They compared how readers process mismatches between fronted temporal adverbials and tense marking on verbs as in (3).

3. a. Last week/*since the summer, James *went* swimming every day. *(Past tense)*
   b. *Last week/since the summer, James *has gone* swimming every day. *(Present perfect)*

In (3a), the temporal adverbial *last week* lexically marks a temporal event with reference and event time preceding speech time, which requires a matching use of grammatical past tense on the verb (Reichenbach, 1947). In contrast, the adverbial *since the summer* in (3b) refers to a temporal event with current relevance or ongoingness, that is, reference time and speech time are identical, with event time preceding both. These adverbials require the use of the present perfect as a grammatical tense marking. Previous research on monolingual English speakers (Baggio, 2008; Steinhauer & Ullman, 2002) finds that native speakers are sensitive to tense mismatches in sentence processing, which suggests that the sentence-initial adverbials make readers create a prediction to encounter a matching grammatical tense marking on the verb. If this prediction is violated by the occurrence of a mismatching grammatical tense, processing difficulty ensues. In a self-paced reading task, Roberts and Liszka (2013) found that L1 French intermediate to advanced learners of English, yet not L1 German learners, were equally sensitive to mismatches between lexical and grammatical tense marking, with longer reading times on regions following the verb for mismatches than matches. Roberts and Liszka (2013) argued that the realization of aspectual distinctions in the L1 restricts the online use of grammatical tense in the L2. Unlike English and French, German does not have grammatical aspect. In addition, the present perfect in German can refer to either past and completed events or events that have current relevance (Rothstein, 2008).

Experiment 2 examines the extent to which additional massed input that illustrates the target relations between temporal adverbials and grammatical tense leads to grammatical predictions in L1 German learners. We compare changes in processing of the experimental group versus a control group. The experimental group received exposure to sentences illustrating differences in the matching of adverbials to past tense and present perfect, respectively (4a). In contrast, the control group read sentences illustrating the matching of adverbials to past tense only (4b). We predict that the processing patterns will change for past tense verbs as well as for present perfect verbs for the experimental group, as the adverbials create predictions for grammatical tense marking, yet not for the control group. Instead, the control group may show adaptation with present perfect verbs only, since the adverbials used in the exposure phase create predictions for past tense verbs (3a), which should lead to slowdowns for mismatched present perfect verbs in (3b).

### *Participants*

Sixty-seven L1 German intermediate late L2 learners took part in the study (44 female, age 21–38 years, mean age: 24.7 years). All participants had started learning English later than age 5, and they were students of subjects other than English at a German university. In line with Roberts and Liszka (2013) we aimed for less proficient learners in Experiment 2 compared to Experiment 1, as the contingencies between temporal adverbials and verb tense are subject to instruction. The LexTALE task (Lemhöfer & Broersma, 2012) placed them into the intermediate range. They participated voluntarily for compensation of 7.50 Euros. Additional participant information is given in Table 5. One-way ANOVAs with the between-subject factor group did not show any significant differences in any measure (all $ps > .08$).

**Table 5.** Experiment 2: Participant information, means (standard deviation)—all participants ($n = 67$)

| | Experimental group | Control group |
|---|---|---|
| *N* | 32 [35] | 32 |
| LexTale score | 69.8 (11.2) | 71.1 (10.5) |
| Length of exposure (yr) | 15.6 (2.5) | 16.3 (3.5) |
| Age of onset (yr) | 9.1 (1.5) | 8.3 (2.0) |

## *Methods and materials*

### *Reading study*

For the pretest and postest, the materials were partially adopted from Roberts and Liszka (2013), and 24 comparable sets of sentences as in (3) were designed (see online-only Supplement for all materials). In addition, 120 filler sentences were constructed that did not contain sentence-initial temporal adverbials and used different tenses. All experimental sentences and half of the fillers were followed by comprehension questions.

The 48 experimental items were split into two sets for which four lists were created according to a Latin square design. The sets were assigned to the pretest and the posttest according to the same procedure as in Experiment 1.

### *Exposure phase*

For the exposure phase, 24 pairs of sentences were constructed as in (4).

4. a. Last year, Bruno liked the new classmate, but he has since then made other friends than him. Now they do not even talk to each other. *(Experimental group)*
   b. Last year, Bruno liked the new classmate, but he did back then not make a new friend in him. Now they do not even talk to each other. *(Control group)*

The experimental group read sentences as in (4a), which displayed a sequence of tense (past-perfect-present). The first sentence contained a temporal adverbial (e.g., *last year*) that was matched with past tense, and the following temporal adverbial (e.g., *since then*) was paired with the matching present perfect. Finally, a sentence paired an adverbial like *now* with present tense. The control group read sentences as in (4b), in which two adverbials (e.g., *last year* and *back then*), were paired with past tense verbs, and a subsequent sentence in present tense (past-past-present). Again, matching adverbials accompanied the tensed verbs; yet, critically, there was no contrast between the use of past tense and present perfect. The 24 filler items that were added contained present tense and future verbs only.

### *Cloze Test and acceptability judgment task*

In order to test whether participants have explicit knowledge of the use of English tenses, we administered a cloze test and an acceptability judgment task, modeled on tasks in Roberts and Liszka (2013). In the cloze test, participants read 30 sentences containing missing verbs, 10 in past tense, 10 in present perfect, and 10 in present tense. Participants added inflected verbs in the gaps, using verbs provided in infinitival form in parentheses. For the acceptability judgment task, participants received 12 experimental items from the pretest, 6 grammatical and 6 ungrammatical. Participants rated the acceptability of these sentences plus 12 fillers by making a binary decision. Both tasks were administered after the reading studies so as to avoid any carryover effects of the offline to the online tasks.

### *Procedure*

The sequence of tasks was identical as in Experiment 1. In addition, the participants completed the cloze test and acceptability judgment task at the end of the session. The self-paced reading experiments were administered using E-Prime 2.0 (Schneider, Eschman, & Zuccolotto, 2002) on a laptop. The sentences were presented in a Moving Windows word-by-word format in 18-point Courier New in white on black on a 15-inch TFT screen. In all, the participants took between 12 and 20 min to complete the pretest and between 18 and 32 min to complete the combined exposure phase and posttest.

### *Analysis and results*

Three participants were excluded from analysis, because they did not complete all tasks ($n = 2$) or had excessively long reading times ($n = 1$). For all of the remaining 64 participants, reading times on the individual segments were trimmed for outliers in that reaction times that were more than 2 *SD* from a participant's mean in this condition and test were trimmed to the mean plus 2 *SD*. In all, this affected less than 4% of the data. Following Roberts and Liszka (2013), we defined four regions of interest in the experimental sentences: the verb (i.e., the main verb in the past tense conditions and the auxiliary in the present perfect conditions), and the three following words each. We analyzed the reading times for the past tense and the present perfect conditions separately, because the segments contained different words. Analyses of reading time measures were conducted using mixed-effect regression models with the lme4 and lmerTest packages in *R* version 3.5.1 (R Development Core Team, 2018). To ensure comparability with the study by Roberts and Liszka (2013), we analyzed the raw reading times.[3] We entered Match (mismatch vs. match), Test (pretest vs. posttest), and Group (experimental vs. control) as fixed factors, including their interactions. We also used Proficiency (LexTALE score) as a continuous and centered fixed effect as an interaction term. The random effects structure was as in Experiment 1.

We present the findings from the offline tasks first, although these tasks were administered after the reading tasks. In both the cloze test and the acceptability judgment task, the groups did not demonstrate any significant differences (Table 6). As these tasks were administered after the reading experiments, the findings suggest that between-group differences in the exposure phase did not affect explicit knowledge of tense. A repeated-measures ANOVA for the acceptability judgment task with the within-subject factors tense and match and the between-subject factor group did not show any effect of Group, $F\,(1, 62) = 0.131$; $p = .719$, yet a significant interaction between Tense and Match, $F\,(1, 62) = 6.608$; $p = .013$,

**Table 6.** Experiment 2: Results of off-line tasks by group ($n = 64$)—standard deviations in parentheses

| | Experimental group ($n = 32$) | Control group ($n = 32$) |
|---|---|---|
| Cloze test – Past tense (in %) | 89.7 (14.9) | 91.9 (12.6) |
| Cloze test – Present perfect | 12.8 (27.5) | 15.9 (25.0) |
| AJT – Past tense (in %) | 53.1 (13.0) | 53.1 (14.9) |
| Past tense match | 94.8 (12.3) | 92.7 (18.4) |
| Past tense mismatch | 11.5 (25.3) | 13.5 (23.7) |
| AJT – Present perfect (in %) | 51.6 (16.6) | 49.5 (16.7) |
| Present perfect match | 85.4 (20.6) | 84.4 (25.4) |
| Present perfect mismatch | 17.7 (26.7) | 14.6 (26.7) |

AJT - Acceptability Judgement Task.

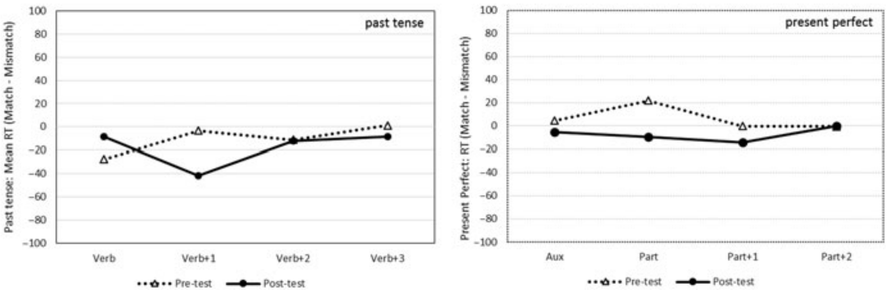

**Figure 3.** Experiment 2: Mean reading time differences (in ms) between match and mismatch condition by verb tense and segment. Experimental group ($n = 32$).

which indicates that mismatches in the present perfect were judged more accurately than in the past tense.

Next, we present the findings from the reading study. Comprehension accuracy was uniformly high at 92.6% ($SD = 6\%$) in the pretest and 92.6% ($SD = 7\%$) in the posttest. The groups did not differ in any test (all $ps > .4$). Figures 3 and 4 illustrate the reading time differences between the match condition minus the mismatch condition for either tense in the pretest and the posttest by group. The match conditions involve a verb in past tense or present perfect paired with a matching adverbial, for example, *last year* or *since last year*, respectively. The mismatch conditions refer to sentences containing a verb in past tense or present perfect paired with a mismatching adverbial. Negative reading times denote sensitivity to tense mismatches, because the mismatch condition was read more slowly than the match condition.

For the experimental group in Figure 3, the reading time differences for past tense verbs change from pretest to posttest in the verb+1 segment. For present perfect verbs, reading profiles are flat. For the control group, Figure 4 shows flat reading times around the zero mark. There is no change from pretest to posttest. Table 7 lists the reading times for the respective groups in the pretest and the posttest.

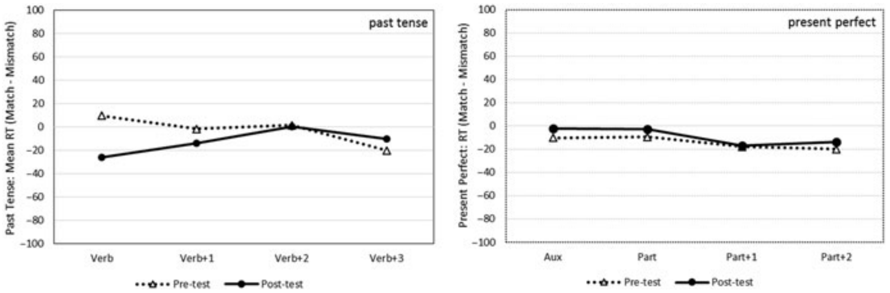

**Figure 4.** Experiment 2: Mean reading time differences (in ms) between match and mismatch condition by verb tense and segment. Control group ($n = 32$).

For past tense verbs, the models returned main effects of test on all segments, showing that participants read faster in the posttest. In addition, there was a marginal interaction between Group, Match, and Test for the verb region ($\beta = 57.50$, $SE = 33.91$, $t = 1.696$, $p = .091$) as well as a marginal interaction between Match and Test for the verb+1 region ($\beta = 38.57$, $SE = 22.15$, $t = 1.742$, $p = .082$). For present perfect verbs, the models returned main effects of Test and a significant interaction of Group, Match, Test, and Proficiency on the participle region ($\beta = 64.66$, $SE = 30.87$, $t = 1.997$, $p = .046$). In addition, for the part+2 region, there was a significant interaction of Match, Group, and Proficiency ($\beta = 46.37$, $SE = 21.91$, $t = 2.116$, $p = .035$).

In view of the interactions with Group and Test, we next report models by group and test. To test for adaptation effects, we checked for interactions of Match and Test. For past tense verbs, the control group demonstrated main effects of Test on all segments, yet no effects of Match or any interactions of Match and Test. For the experimental group, beyond significant main effects of Test on all segments, there was a significant main effect of Match in the verb+1 region ($\beta = 41.91$, $SE = 12.51$, $t = 3.350$, $p < .001$), which was qualified by a Match × Test interaction ($\beta = 38.56$, $SE = 17.69$, $t = 2.180$, $p = .030$). Follow-up analyses by test showed no significant effect of Match for the pretest ($\beta = 3.44$, $SE = 13.94$, $t = 0.247$, $p = .805$, $d = 0.08$), yet a significant effect of Match in the posttest ($\beta = 41.73$, $SE = 10.50$, $t = 3.976$, $p < .001$, $d = 0.31$). Proficiency did not have any effect.

For present perfect verbs, there were no effects or interactions with Match for the experimental group on any segment. For the control group, a marginal interaction of Match and Proficiency obtained on the part+2 region ($\beta = 32.85$, $SE = 16.93$, $t = 1.940$, $p = .053$). Follow-up analyses according to a median split yielded no significant effects of Match in either the pretest or the posttest.

In sum, whereas neither group showed reading-time differences for present perfect verbs, the experimental group demonstrated a significant effect of Match in the posttest for the verb+1 region for past tense verbs, showing a slowdown for past tense verbs paired with a temporal adverbial denoting current relevance.

### *Discussion of Experiment 2*

Experiment 2 investigated if exposure to sentences that illustrate the use of grammatical tense in English with matching temporal adverbials leads L2 learners to become sensitive to mismatches between temporal adverbials and grammatical tense inflection.

**Table 7.** Experiment 2: Reading times (in ms) and standard deviations in parentheses by time, verb tense, match, and group ($n = 32$ each)

| | | Group | Experimental group | | | | Control group | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Tense | Time | Match | Verb | Verb+1 | Verb+2 | Verb+3 | Verb | Verb+1 | Verb+2 | Verb+3 |
| Past | Pretest | Match | 473 (153) | 446 (109) | 465 (143) | 473 (143) | 523 (153) | 495 (140) | 488 (128) | 480 (126) |
| | Pretest | Mismatch | 501 (177) | 449 (119) | 476 (164) | 472 (129) | 514 (170) | 497 (146) | 487 (143) | 501 (165) |
| | Posttest | Match | 393 (112) | 360 (84) | 385 (107) | 395 (111) | 411 (118) | 400 (89) | 395 (93) | 396 (84) |
| | Posttest | Mismatch | 401 (108) | 402 (114) | 397 (111) | 404 (120) | 437 (135) | 414 (126) | 395 (94) | 406 (119) |
| | | | Aux | Part | Part+1 | Part+2 | Aux | Part | Part+1 | Part+2 |
| Present Perfect | Pretest | Match | 449 (133) | 479 (170) | 453 (131) | 462 (136) | 481 (130) | 474 (124) | 455 (108) | 458 (106) |
| | Pretest | Mismatch | 445 (122) | 457 (135) | 453 (112) | 462 (139) | 491 (143) | 484 (134) | 472 (139) | 478 (129) |
| | Posttest | Match | 372 (88) | 388 (116) | 369 (86) | 385 (95) | 409 (95) | 397 (90) | 379 (88) | 395 (94) |
| | Posttest | Mismatch | 377 (94) | 398 (114) | 383 (98) | 385 (105) | 411 (116) | 400 (114) | 396 (96) | 409 (134) |

The pretest results replicate Roberts and Liszka (2013) in that L1 German learners of L2 English were not sensitive to tense mismatches in reading. Having read informative sentences that illustrate the pairing of adverbials and verb tense in both past tense and present perfect, the experimental group came to show slowdowns for past tense mismatches in the posttest. Reading patterns in the control group did not change at all from pre- to posttest. These findings confirm the hypothesis regarding group differences between the experimental and the control group. However, adaptation to the use of grammatical tense marking was selective to past tense mismatches in the experimental group. Temporal adverbials like *since last night* gave rise to a prediction for encountering a present perfect verb. When instead a past tense verb was encountered, reading slowed down. However, the experimental group did not demonstrate reading time differences for present perfect verbs, that is, a temporal adverbial expressing complete past reference (e.g., *last night*) did not create a strong expectation for a past tense verb, with present perfect being equally permissible.

It is likely that the selective adaptation effect reflects L1 properties. In German, combining perfect tense verbs with an adverbial denoting a completed event in the past is grammatical (Rothstein, 2008). It would appear that the exposure could only partially sensitize learners to the constraints on temporal reference in English, and L1 options persevered despite exposure to sentences that illustrate the mapping of tense in the L2. Along these lines, the interactions with proficiency in the present perfect sentences further suggest that target processing of mismatches in the present perfect may only be mastered by more proficient learners.

Of note, the group differences in adaptation to past tense mismatches did not translate into group differences in the cloze test and the acceptability judgment task that tapped into explicit knowledge of tense. In both of these tasks, knowledge of tense mismatches remained low, and the experimental group did not show an advantage over the control group. This lack of transition of adaptation from online to offline tasks may be a matter of degree in that more exposure would have been necessary for offline performance to be affected; alternatively, implicit learning through adaptation in comprehension may not feed into explicit knowledge (for discussion, see Ellis et al., 2009). In all, the findings from Experiment 2 demonstrate that learners can overcome L1-induced insensitivity to grammatical tense inflection by adapting their processing after exposure to informative input.

## GENERAL DISCUSSION

The present paper tested whether inflectional variability in adult L2 learners reduces as a consequence of morphosyntactic adaptation to target input in two processing experiments. Both experiments found interactions with Group bearing out that the experimental group adapted to target processing patterns following systematic and massed exposure to target input. In contrast, the control group, which received exposure to comparable input that lacked systematic information about the target structure, did not show adaptation.

In both experiments, learners in the experimental groups demonstrated selective adaptation. In Experiment 1, adaptation was specific to case marking, and the experimental group did not show attenuation of garden paths occasioned by the intransitive use of an optionally transitive verb across all conditions. In Experiment 2, adaptation was particular to past tense mismatches, and the experimental group continued to be insensitive to present perfect mismatches.

In both experiments, the results suggest that the parser extracted relevant information from the input provided in the exposure phase to adjust its processing in the posttest. As the input in the exposure phase did not duplicate the experimental materials, it is unlikely that the targetlike processing of the experimental group in the posttest reflects increasing practice with the sentences or low-level frequency-driven learning. The lack of interactions with trial number in Experiment 1 suggests that merely reading a number of experimental sentences did not lead to adaptation in the L2 groups (see Kaan et al., 2019). Instead, more targetlike processing resulted from the exposure phase which was designed to elicit prediction errors and flag correct parses in the experimental group. The experimental sentences in the exposure phase of Experiment 1 were designed for the parser to systematically experience prediction error in that a potentially transitive verb was consistently not followed by an object. Subsequently, the parser used the information that became available, that is, nominative case marking on the pronoun, as evidence that clearly pointed to the target structure (e.g., Fodor & Inoue, 2000; Hopp, 2006). In the absence of an informative cue flagging prediction error and triggering reanalysis, the parser continued to be garden-pathed, as can be seen in persistent garden paths for the implausible and plausible sentences (1d and 1e) in the posttest.

In Experiment 2, the parser appeared to become sensitive to detecting tense mismatches in the past-tense condition after massed input illustrating relations between temporal adverbials and tensed verbs. Once the experimental group had read the sentences in the exposure phase, the group predicted an adverbial like *since last night* to be followed by a present perfect verb; when it was followed by a past tense verb instead, readers slowed down. Critically, the individual pairings of a past adverbial and past tense marking on verbs was apparently not sufficient to sensitize the parser to these mappings. The control group received twice as many tokens of past tense matches as the experimental group in the exposure phase; however, only the experimental group became sensitive to past tense mismatches in the posttest. In other words, the simple input frequency of matching adverbial-verb combinations in past tense did not lead to target processing. Rather, it appears that the systematic and discriminating use of tenses in the input provided to the experimental group, that is, past tense versus present perfect, was a prerequisite for the parser to come to predict relations between temporal adverbials and verb tense, and, in turn, to detect past tense mismatches. In this regard, it was somewhat unexpected that the experimental group did not come to make reading time differences in the present perfect condition. We speculated that L1 mappings constrain adaptation. In these cases, more exposure or higher proficiency, that is, more experience with target input on part of the learners may be required for adaptation to occur. We leave these issues for further research.

Irrespective of whether prediction and prediction error are the specific mechanisms underlying the emergence of target processing in the experimental groups, the present study extends previous findings on syntactic adaptation among native speakers and L2 learners to the processing of inflection in sentence contexts. Crucially, L2 learners do not only adapt their syntactic processing preferences (e.g., Arai, 2016; Beatty-Martínez & Dussias, 2018); rather, the present study shows that L2 learners come to integrate inflection incrementally in L2 parsing as a function of massed short-term exposure to target input. This study finds adaptation across two different experimental methods, self-paced reading and eye tracking during reading, and two phenomena, suggesting that L2 morphosyntactic adaptation is not method specific (see Yan, Farmer, & Jaeger, 2019) or particular to one type of phenomenon. Future studies should aim to replicate these findings across different tasks, other learner groups and extend it to other phenomena.

Finding such rapid L2 adaptation to inflection conflicts with the results on L2 syntactic adaptation in Kaan et al. (2019), in which L2 speaker did not adapt to filled-gap constructions and clausal coordination. Even though there are many differences between the studies in design and in the number of items, a crucial difference is that studies on syntactic adaptation investigate shifting processing preferences between two parses of an ambiguous structure that differ in frequency. In contrast, the present study tested adaptation to unambiguous target input by upping its frequency through massed input. Clustered or blocked input as provided by the exposure phase has been shown to effect faster and stronger adaptation than the same number of exposures that are distributed more widely (see Myslín & Levy, 2016). Such extra blocked and unambiguous input adds to the experience learners have with the target language in general, and it may provide critical tipping points for the parser to integrate inflection. In any case, it will be fruitful to investigate potentially different signatures and time courses of adapting to ambiguities versus adapting to the target use of inflection.

From an L2 perspective, the crucial issue, though, is whether adaptation in the two experiments reflects implicit learning of inflection. According to error-based implicit learning models, language users generate predictions about the input and adjust these if they experience prediction errors (e.g., Dell & Chang, 2014). According to belief-updating models (e.g., Jaeger & Snider, 2013), learners adjust their probability computations as to how likely a structure is to occur if the input does not match the priors. In both approaches, language users adapt to specific structures and will continuously adjust their expectations if the input changes. In several respects, these notions of implicit learning are different from the idea of implicit learning espoused in approaches to L2 acquisition, which conceptualize it as leading to relatively stable representations or knowledge that learners can call upon in production and comprehension (e.g., Hulstijn, 2005). First, adaptation is dynamic and malleable in that changes in the input can rapidly modulate expectations, sometimes across a small number of trials (e.g., Henry et al., 2017; Hopp, 2016). Second, adaptation effects are rather specific to the structures tested in the experiments. Since users take multiple cues into account when predicting, for example, speaker identity (Kamide, 2012), contextual and lexical information (Ryskin, Qi, Duff, & Brown-Schmidt, 2017), the scope of adaptation to the target use of inflection in the two experiments may be limited to the particular sentence contexts in which the participants encountered case and tense marking, respectively. As a case in point, the adaptation in reading times in the experimental group to past tense mismatches in Experiment 2 did not translate into an advantage of the experimental over the control group in the cloze test and acceptability task. The lack of offline advantages may be due to differences in tasks (reading vs. judgments) or sentence contexts (reading vs. cloze test). In these respects, future studies should investigate possible transitions of online adaptation in processing to offline performance in explicit tasks.

What's more, Experiment 1 used the same verbs in the testing and the exposure sessions. Although adaptation in Experiment 1 reflected the interaction of verb transitivity and the use of nominative case marking, it may be argued that the adaptation effects partially index lexical rather than morphosyntactic adaptation (Kaan & Chun, 2018b; Ryskin et al., 2017). Future research should gauge the degree of lexical abstraction of morphosyntactic adaptation and its potential interactions.

Moreover, we need research on how morphosyntactic adaptation interacts with L2 proficiency. On the one hand, lower proficiency learners with less exposure to the target language may have less entrenched nontarget L2 processing, experience greater prediction error, and thus adapt faster than higher proficiency learners when exposed to massed target input; on the other hand, higher proficiency learners may be more targetlike in L2 processing to begin with, so that they require less additional exposure to adapt.

In addition, the study found adaptation in a pre-/posttest design within a single lab session. Therefore, it cannot speak to the extent to which adaptation effects persevere in the long run and whether they generalize beyond the lab context. Since language users adapt to context (Kaschak, Kutta, & Coyle, 2014) and speaker-specific input distributions (e.g., Kroczek & Gunter, 2017), generalization of adaptation observed in lab experiments beyond the testing situation cannot be taken for granted and needs to be tested. Finally, the study was restricted to L1 German learners of English, and it will be useful to explore adaptation in cross-linguistic comparisons to investigate how L1–L2 differences interact with adaptation.

Despite these many open issues, the study nevertheless yields some initial evidence to show that morphosyntactic adaptation attenuates inflectional variability in L2 learners by occasioning targetlike processing following massed exposure. Finding such morphosyntactic adaptation among adult L2 learners has implications for research on L2 processing.

First, rapid adaptation to target use of inflection in L2 processing casts doubt on claims that processing data afford direct insights into grammatical representation of L2 learners (e.g., Clahsen & Felser, 2006). As (L2) processing is adaptive to short-term changes in the input, parsing data do not allow straightforward inferences about grammatical representations. In particular, approaches to inflectional variability in L2 acquisition that partially rely on processing data to argue that L2 learners lack morphosyntactic representations (representational deficit approaches; e.g., Hawkins & Chan, 1997; Tsimpli & Dimitrakopoulou, 2007) need to accommodate effects of adaptation, that is, changes in sensitivity to inflection. If anything, finding that L2 learners can adapt to target uses of inflection in L2 processing presupposes or entails some form of grammatical representation according to formal approaches to L2 acquisition.

Second, adaptation opens a new perspective on comparisons between native and nonnative processing. So far, L2 processing has investigated differences in the types or timing of grammatical knowledge applied in real-time sentence comprehension by L2 learners (Clahsen & Felser, 2018; Cunnings, 2017). In addition, it will be interesting to examine whether L1 and L2 processors adapt similarly to changes in the use of grammatical information in the input. Such a perspective inherently entails a shift in focus from one-time snapshots of L2 processing toward L2 development and can thus link up L2 processing research more closely with research on L2 acquisition.

In this vein, studies on adaptation in L2 sentence processing may also pave the way toward cross-disciplinary exchanges between research in L2 sentence processing, on the one hand, and studies on incidental learning in L2 acquisition (e.g., Rebuschat & Williams, 2012) as well as research on differences between implicit versus explicit learning (e.g., Ellis et al., 2009), on the other hand. In conclusion, the study of morphosyntactic adaptation holds promise to link processing directly to acquisition and learning. Future research should gauge the scope of adaptation in L2 processing across different phenomena and its potential to account for L2 development and possible fossilization in processing terms.

**Supplementary material.** To view supplementary material for this article, please visit https://doi.org/10.1017/S0142716420000119

**Acknowledgements.** I would like to thank Tina Matthias for help with stimulus creation and data collection for Experiment 2 as well as Lina Hanisch and Leah Wildner for help with data collection for Experiment 1. I am grateful to the audiences at GASLA 15 in Reno, Nevada, SLRF 2019 at Michigan State as well as the CLS colloquium at Penn State University for comments and questions. Finally, I wish to thank Carrie N. Jackson, the anonymous reviewers and the Associate Editor at *Applied Psycholinguistics* for helpful comments that led to improvements of the manuscript.

## NOTES

**1.** Note that the pronouns in the exposure phase were deliberately designed to be coreferential with a previous noun phrase in the sentence. This way, they differed from the pronouns in the pre- and posttest that were not coreferential with the subject of the adjunct clause. Any potential adaptation effects in the posttest could thus not result from participants getting used to the nonanaphoric use of personal pronouns. Moreover, the input to both the experimental and the control group was matched in terms of the anaphoricity of the pronouns, so that any between-group differences in the pretest must reflect the formal properties of the pronouns, for example, case marking, rather than their interpretive properties.

**2.** Following Kaan et al. (2019), we also ran the models on log-transformed reading times (natural log), and we obtained comparable adaptation effects between pretest and posttest that were limited to the case comparison and the experimental group. In addition, the critical interaction between Test and Group for the pronoun condition also became significant ($\beta = 0.18$, $SE = 0.08$, $t = 2.375$, $p = .018$), underscoring that the group differences remain when the skewedness of the data distribution is adjusted for.

**3.** The models for log-transformed reading times showed similar effects for past tense verbs, that is, they yielded a marginal interaction of Group and Test ($\beta = 0.07$, $SE = 0.04$, $t = 1.821$, $p = .069$) for the verb region and, for the verb+1 region, significant interactions of Group and Match ($\beta = 0.08$, $SE = 0.04$, $t = 2.051$, $p = .041$) and Match and Test ($\beta = 0.08$, $SE = 0.04$, $t = 2.178$, $p = .029$). For present perfect verbs, the interaction of Match, Group and Proficiency on the verb+3 region became marginally significant ($\beta = 0.07$, $SE = 0.04$, $t = 1.740$, $p = .082$). For the experimental group, the critical Match by Tense interaction for past tense verbs in the verb+1 region was significant ($\beta = 0.08$, $SE = 0.04$, $t = 2.234$, $p = .026$).

## References

**Andrews, C., Dillon, B., & Staub, A.** (2019). *Fool me once: Readers Adapt to NP/Z garden paths but not ORCs.* Poster presented at the 32nd Annual CUNY Conference on Human Sentence Processing, Boulder, CO, March 2019.

**Arai, M.** (2016). *Importance of making (prediction) errors in L2 processing.* Poster presented at the 22nd Annual Conference on Architectures and Mechanisms for Language Processing, Bilbao, Spain, September 2016.

**Baggio, G.** (2008). Processing temporal constraints: An ERP study. *Language Learning*, **58**, 35–55.

**Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J.** (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory & Language*, **68**, 255–278.

**Beatty-Martínez, A. L., & Dussias, P. E.** (2018). Tuning to languages: Experience-based approaches to the language science of bilingualism. *Linguistics Vanguard*, **2018**, 20170034

**Christiansen, M. H., & Chater, N.** (2016). The Now-or-Never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, **39**, e62.

**Clahsen, H., & Felser, C.** (2006). Grammatical processing in language learners. *Applied Psycholinguistics*, **27**, 3–42.

**Clahsen, H., & Felser, C.** (2018). Some notes on the shallow structure hypothesis. *Studies in Second Language Acquisition*, **40**, 693–706.

**Cunnings, I.** (2017). Parsing and working memory in bilingual sentence processing. *Bilingualism: Language and Cognition*, **20**, 659–678.

**Dell, G. S., & Chang, F.** (2014). The P-chain: Relating sentence production and its disorders to comprehension and acquisition. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **369**, 1–9.

**Dussias, P. E., Valdés Kroff, J. R., Guzzardo Tamargo, R. E., & Gerfen, C.** (2013). When gender and looking go hand in hand: Grammatical gender processing in L2 Spanish. *Studies in Second Language Acquisition*, **35**, 353–387.

**Ellis, R., Loewen, S., Elder, C., Reinders, H., Erlan, R., & Philip, J.** (2009). *Implicit and explicit knowledge in second language learning, testing and teaching* (Vol. 42). Bristol: Multilingual Matters.

**Fine, A. B., Jaeger, T. F., Farmer, T. A., & Qian, T.** (2013). Rapid expectation adaptation during syntactic comprehension. *PLOS ONE*, **8**, 1e18.

**Fodor, J. D., & Inoue, A.** (2000). Syntactic features in reanalysis: Positive and negative symptoms. *Journal of Psycholinguistic Research*, **29**, 25–36.

**Gambi, C., Pickering, M. J., & Rabagliati, H.** (2016). Beyond associations: Sensitivity to structure in preschoolers' linguistic predictions. *Cognition*, **157**, 340–351.

**Grüter, T., Rohde, H., & Schafer, A. J.** (2017). Coreference and discourse coherence in L2. *Linguistic Approaches to Bilingualism*, **7**, 199–229.

**Harrington Stack, C., James, A., & Watson, D. G.** (2018). A failure to replicate rapid syntactic adaptation in comprehension. *Memory and Cognition*, **46**, 864–877.

**Hartsuiker, R. J., Beerts, S., Loncke, M., Desmet, T., & Bernolet, S.** (2016). Cross-linguistic structural priming in multilinguals: Further evidence for shared syntax. *Journal of Memory and Language*, **90**, 14–30.

**Havron, N., de Carvalho, A., Fiévet, A. C., & Christophe, A.** (2019). Three- to four-year-old children rapidly adapt their predictions and use them to learn novel word meanings. *Child Development*, **90**, 82–90.

**Hawkins, R., & Chan, C. Y. H.** (1997). The partial availability of Universal Grammar in second language acquisition: The "failed functional features hypothesis." *Second Language Research*, **13**, 187–226.

**Henry, N., Hopp, H., & Jackson, C. N.** (2017). Cue additivity and adaptivity in predictive processing. *Language, Cognition and Neuroscience*, **32**, 1229–1249.

**Hopp, H.** (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research*, **22**, 369–397.

**Hopp, H.** (2010). Ultimate attainment in L2 inflectional morphology: Performance similarities between non-native and native speakers. *Lingua*, **120**, 901–931.

**Hopp, H.** (2013). Grammatical gender in adult L2 acquisition: Relations between lexical and syntactic variability. *Second Language Research*, **29**, 33–56.

**Hopp, H.** (2014). Individual differences in the L2 processing of object-subject ambiguities. *Applied Psycholinguistics*, **36**, 129–173.

**Hopp, H.** (2015). Semantics and morphosyntax in predictive L2 sentence processing. *International Review of Applied Linguistics in Language Teaching*, **53**, 277–306.

**Hopp, H.** (2016). Learning (not) to predict: Grammatical gender processing in second language acquisition. *Second Language Research*, **32**, 277–307.

**Hoshino, N., Dussias, P. E., & Kroll, J. F.** (2010). Processing subject–verb agreement in a second language depends on proficiency. *Bilingualism: Language and Cognition*, **13**, 87–98.

**Huettig, F., & Mani, N.** (2016). Is prediction necessary to understand language? Probably not. *Language, Cognition and Neuroscience*, **31**, 19–31.

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

**Ito, A., Pickering, M. J., & Corley, M.** (2018). Investigating the time-course of phonological prediction in native and non-native speakers of English: A visual world eye-tracking study. *Journal of Memory and Language*, **98**, 1–11.

**Jackson, C. N., & Dussias, P. E.** (2009). Cross-linguistic differences and their impact on L2 sentence processing. *Bilingualism: Language and Cognition*, **12**, 65–82.

**Jackson, C. N., & Ruf, H. T.** (2016). The priming of word order in second language German. *Applied Psycholinguistics*, **38**, 315–345.

**Jaeger, T. F., & Snider, N. E.** (2013). Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, **127**, 57–83.

**Jiang, N.** (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, **57**, 1–33.

**Kaan, E.** (2014). Predictive sentence processing in L2 and L1: What is different? *Linguistic Approaches to Bilingualism*, **4**, 257–282.

**Kaan, E., & Chun, E.** (2018a). Priming and adaptation in native speakers and second-language learners. *Bilingualism: Language and Cognition*, **21**, 228−242.

**Kaan, E., & Chun, E.** (2018b). Syntactic adaptation. *Current Topics in Language*, **68**, 85–116.

**Kaan, E., Futch, C., Fuertes, R. F., Mujcinovic, S., & de la Fuente, E. Á.** (2019). Adaptation to syntactic structures in native and nonnative sentence comprehension. *Applied Psycholinguistics*, **40**, 3–27.

**Kamide, Y., Scheepers, C., & Altmann, G. T.** (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of psycholinguistic research*, **32**, 37–55.

**Kamide, Y.** (2012). Learning individual talkers' structural preferences. *Cognition*, **124**, 66–71.

**Kaschak, M. P., Kutta, T. J., & Coyle, J. M.** (2014). Long and short term cumulative structural priming effects. *Language, Cognition and Neuroscience*, **29**, 728–743.

**Kleinschmidt, D. F., & Jaeger, T. F.** (2015). Robust speech perception: Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, **122**, 148–203.

**Kroczek, L. O. H., & Gunter, T. C.** (2017). Communicative predictions can overrule linguistic priors. *Scientific Reports*, **7**, 17581.

**Kuperberg, G. R., & Jaeger, T. F.** (2016). What do we mean by prediction in language comprehension? *Language, Cognition and Neuroscience*, **31**, 32–59.

**Lemhöfer, K., & Broersma, M.** (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavioral Research*, **44**, 325–343.

**Lew-Williams, C., & Fernald, A.** (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science*, **18**, 193–198.

**Lew-Williams, C., & Fernald, A.** (2010). Real-time processing of gender-marked articles by native and non-native Spanish speakers. *Journal of Memory and Language*, **63**, 447–464.

**Lukyanenko, C., & Fisher, C.** (2016). Where are the cookies? Two- and three-year-olds use number-marked verbs to anticipate upcoming nouns. *Cognition*, **146**, 349–370.

**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.

**Mani, N., & Huettig, F.** (2012). Prediction during language processing is a piece of cake—But only for skilled producers. *Journal of Experimental Psychology: Human Perception and Performance*, **38**, 843.

**Mitsugi, S., & MacWhinney, B.** (2016). The use of case marking for predictive processing in second language Japanese. *Bilingualism: Language and Cognition*, **19**, 19–35.

**Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. T.** (2010). Second language acquisition of gender agreement in explicit and implicit training conditions: An event-related potential study. *Language Learning*, **60**, 154–193.

**Myslín, M., & Levy, R.** (2016). Comprehension priming as rational expectation for repetition: Evidence from syntactic processing. *Cognition*, **147**, 29–56.

**Phillips, C., & Ehrenhofer, L.** (2015). The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism*, **5**, 409–453.

**Pickering, M. J., & Gambi, C.** (2018). Predicting while comprehending language: A theory and review. *Psychological Bulletin*, **144**, 1002–1044.

**Pickering, M. J., & Garrod, S.** (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences*, **36**, 329–347.

**Rabagliati, H., Gambi, C., & Pickering, M. J.** (2015). Learning to predict or predicting to learn? *Language, Cognition and Neuroscience*, **31**, 94–105.

**Rayner, K.** (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, **124**, 372–422.

**R Core Team**. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org/.

**Rebuschat, P., & Williams, J. N.** (Eds.) (2012). *Statistical learning and language acquisition* (Vol. 1). Berlin: de Gruyter.

**Reichenbach, H.** (1947). The tenses of verbs. *Time: From concept to narrative construct: A reader*. Göttingen: de Gruyter.

**Roberts, L., & Liszka, S. A.** (2013). Processing tense/aspect-agreement violations on-line in the second language: A self-paced reading study with French and German L2 learners of English. *Second Language Research*, **29**, 413–439.

**Rothstein, B.** (2008). *The perfect time span: On the present perfect in German, Swedish and English* (Vol. 125). Amsterdam: Benjamins.

**Ryskin, R. A., Qi, Z., Duff, M. C., & Brown-Schmidt, S.** (2017). Verb biases are shaped through lifelong learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **43**, 781.

**Schneider, W., Eschman, A., & Zuccolotto, A.** (2002). *E-Prime reference guide*. Pittsburgh, PA: Psychology Software Tools.

**Staub, A.** (2007). The parser doesn't ignore intransitivity, after all. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **33**, 550–569.

**Staub, A., & Rayner, K.** (2007). Eye movements and on-line comprehension processes. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 327–342). Oxford: Oxford University Press.

**Steinhauer, K., & Ullman, M. T**. (2002). Consecutive ERP effects of morpho-phonology and morpho-syntax. *Brain and Language*, **83**, 62–65.

**Traxler, M. J.** (2002). Plausibility and subcategorization preference in children's processing of temporarily ambiguous sentences: Evidence from self-paced reading. *The Quarterly Journal of Experimental Psychology: Section A*, **55**, 75–96.

**Traxler, M. J., & Pickering, M. J.** (1996). Case marking in the parsing of complement sentences: Evidence from eye movements. *Quarterly Journal of Experimental Psychology, 49A*, 991–1004.

**Tsimpli, I. M., & Dimitrakopoulou, M.** (2007). The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research*, **23**, 215–242.

**van Heugten, M., & Shi, R.** (2009). French-learning toddlers use gender information on determiners during word recognition. *Developmental Science*, **12**, 419–425.

**Westfall, J., Kenny, D. A., & Judd, C. M.** (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, **143**, 2020–2045.

**Wicha, N. Y., Moreno, E. M., & Kutas, M.** (2004). Anticipating words and their gender: An event-related brain potential study of semantic integration, gender expectancy, and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, **16**, 1272–1288.

**Yan, S., Farmer, T. A., & Jaeger, T. F.** (2019). *Syntactic adaptation in natural reading*. Unpublished manuscript, University of Rochester.

**Cite this article:** Hopp, H. (2020). Morphosyntactic adaptation in adult L2 processing: Exposure and the processing of case and tense violations. *Applied Psycholinguistics* **41**, 627–656. https://doi.org/10.1017/S0142716420000119