CrossMark

Applied Psycholinguistics **36** (2015), 129–173
doi:10.1017/S0142716413000180

Individual differences in the second language processing of object–subject ambiguities

HOLGER HOPP
University of Mannheim

Received: December 5, 2011 Accepted for publication: November 7, 2012

ADDRESS FOR CORRESPONDENCE
Holger Hopp, Department of English Linguistics, University of Mannheim, Schloss EW 266, 68131 Mannheim, Germany. E-mail: hhopp@rumms.uni-mannheim.de

ABSTRACT
This study investigates whether and how individual differences modulate the adult second language (L2) processing of syntactic ambiguities. In a linear mixed regression analysis, we test how proficiency, working memory, reading speed, automaticity in lexical access, and grammatical integration ability affect the resolution of temporary object–subject ambiguities in L2 English. The results from 75 first language German advanced learners attest that individual differences in syntactic integration ability modulate the reliance on morphosyntactic and plausibility information. Similar to native speakers, L2 learners are found to adopt two different routes in L2 processing. The findings highlight the role of individual differences and qualify previous generalizations about the relative use of morphosyntactic and other types of information in L2 processing.

Research on second language (L2) sentence processing investigates how late L2 learners construct sentence representations in real-time comprehension. What emerges as a finding across studies is that adult L2 learners use different information types to different extents in on-line processing. On the one hand, L2 learners reliably recruit semantic, contextual, and plausibility information even at lower proficiency levels (e.g., Hopp, 2009; Hoshino, Dussias, & Kroll, 2010; Juffs & Harrington, 1996; Roberts & Felser, 2011; Williams, Möbius, & Kim, 2001). On the other hand, L2 learners have difficulty integrating morphosyntactic information, such as inflection or syntactic structure, in processing the L2 (Felser & Roberts, 2007; Jiang, 2004; Marinis, Roberts, Felser, & Clahsen, 2005; but see Omaki & Schulz, 2011). These findings have given rise to the observation that L2 learners underrely on grammatical information in processing, instead capitalizing on lexical–semantic and pragmatic information.

Many approaches have generalized this observation as a trait characterizing adult L2 processing across the board (e.g., Clahsen & Felser, 2006b; Jiang, 2004, 2007; Ullman, 2005). However, we know relatively little as to whether and how

© Cambridge University Press 2013 0142-7164/13 $15.00

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

the relative reliance on morphosyntactic and other types of information in L2 processing is modulated by individual differences. Whereas individual differences have been widely studied in L2 acquisition (e.g., Bowden, Sanz, & Stafford, 2005; Dörnyei, 2009) and L1 sentence processing (e.g., Farmer, Misyak, & Christiansen, in press), L2 processing studies are only beginning to consider individual variation among adult L2 learners as a factor contributing to nontarget processing (e.g., Michael & Gollan, 2005). Nevertheless, even the limited evidence available so far indicates that the L2 processing of morphosyntax and other types of grammatical information interacts with cognitive and capacity-related aspects of language processing.

First, as L2 proficiency rises, learners demonstrate growing incremental use of morphosyntax (Hopp, 2006; Jackson & Bobb, 2009; McLaughlin et al., 2010; Rossi, Gugler, Friederici, & Hahne, 2006; Steinhauer, White, & Drury, 2009) as well as target processing strategies (e.g., Frenck-Mestre, 2002). Second, differences in verbal working memory capacity partially affect L2 morphosyntactic processing (Havik, Roberts, van Hout, Schreuder, & Haverkort, 2009; Jackson & Bobb, 2009). Third, lower levels of lexical automaticity in word segmentation and recognition (e.g., Coderre, van Heuven, & Conklin, 2011; Liu & Perfetti, 2003) affect adult L2 performance in grammaticality judgment tasks on morphosyntax (McDonald, 2006) and accuracy in the production of past tense (McDonald & Roussel, 2010). Fourth, adult L2 learners have lower capacities than natives to integrate different types of grammatical information in context (Kilborn, 1992) and therefore may need to variably allocate their resources to different types of information in L2 processing (Hopp, 2009; Sorace, 2005, 2011).

Various individual learner factors hence engender differences in the L2 processing of morphosyntax. However, the picture regarding the contributions of individual factors in L2 processing is as yet rather scattered, because only a handful of studies have considered particular individual differences with respect to one particular grammatical phenomenon.

The aim of the present study is to systematically investigate the relative contributions of various measures of individual differences to L2 processing. To this end, 75 high-intermediate to near-native L1 German speakers of English and 18 native controls were tested on a battery of tasks, encompassing different domains of linguistic processing, as well as an eye tracking study on object–subject ambiguities. Object–subject ambiguities are well suited for investigating effects of individual differences on L2 processing patterns, because the degrees of lexical–thematic, morphosyntactic, and plausibility information available to the reader for incrementally constructing sentence interpretations can be experimentally manipulated. In a mixed regression analysis, we analyze whether proficiency, reading speed, working memory, automaticity in basic linguistic processing, and information integration ability in grammatical contexts affect processing patterns in the L2. The present experiments thus aim to ascertain the degree to which individual differences modulate L2 processing. These findings will be relevant with a view to establishing whether the generalizations about L2 processing arrived at in previous research hold across individuals or whether they mask important individual differences among L2 learners.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

MONOLINGUAL PROCESSING OF OBJECT–SUBJECT AMBIGUITIES

Many studies on native language processing have investigated temporary object–subject ambiguities in garden path sentences like *When Anne bathed the baby fell out of the bathtub* (for an overview, see Staub, 2007). As a consequence of the incremental interpretation of these sentences, readers initially analyze the noun phrase (NP) *the baby* as the direct object of the verb of the preposed adjunct clause *bathe*. When encountering the main verb, however, the initial direct object analysis of the NP turns out to be incorrect, such that readers need to revise their provisional parse. Compared to control sentences with punctuation that do not give rise to such a temporary ambiguity (1b), readers exhibit reading slowdowns on the main clause verb and later regions in (1a), which highlight (a) that they had initially been led down the garden path to adopt a direct object interpretation of the NP and (b) that they reanalyzed this interpretation.

(1) a. When Anne bathed the baby fell out of the bathtub.
b. When Anne bathed, the baby fell out of the bathtub.

In different frameworks of native processing models, this reanalysis process has been conceptualized either in terms of the partial reprocessing of the sentence in serial models of sentence processing (e.g., Mitchell, 1994; Pritchett, 1992; van Gompel, Pickering, Pearson, & Liversedge, 2005) or as a reranking of several analyses maintained in parallel according to constraint-based models (e.g., Spivey & Tanenhaus, 1998). Differences between these models aside, the task for the reader when getting to the main clause verb in (1a) is to undo the thematic association of the NP with the first verb and to establish the NP as the main clause subject by linking it to the main clause verb via thematic role assignment and creating appropriate case marking relations.

It is important that the relative costs of integrating the main clause verb into the current representation of the sentence are modulated by the type of previous information biasing toward an object interpretation of the NP. Hence, the strength of this bias affects the ease of recovery from the initial misanalysis (e.g., Frazier & Clifton, 1998), that is, it determines how quickly readers can undo the association of the NP as the object of the verb of the adjunct clause. Differences in recovery costs have been observed for subcategorization, morphosyntactic, and plausibility information.

First, subcategorization properties of the verb systematically affect reanalysis in that intransitive verbs (e.g., *cried*) give rise to smaller recovery effects than do optionally transitive verbs, such as *bathe* (e.g., Adams, Clifton, & Mitchell, 1998; DeDe 2010; Staub, 2007; Trueswell, Tanenhaus, & Kello, 1993; van Gompel & Pickering, 2001). Second, Traxler (2002) shows that subcategorization preferences of optionally transitive verbs correlate with the degree of reanalysis effort (see also Wilson & Garnsey, 2009). Second, the strength of the reanalysis effect also depends on morphosyntactic information on the noun. Unambiguously nominative marked pronouns (e.g., *she*) in sentences containing reduced complement clause ambiguities like *Dan recognized you/she would be unhappy* prevent readers from getting garden-pathed in the first place (Traxler & Pickering, 1996). On the

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

subsequent main verb, then, no reanalysis costs are visible. Third, the plausibility of the NP as a potential object of the first verb affects the strength of the reanalysis effect (Clifton, 1993). Using eye tracking, Pickering and Traxler (1998) tested sentences like *As the woman edited/sailed the magazine about fishing amused all the reporters*. Whereas the NP *the magazine about fishing* is a plausible object of the verb *edited*, it does not serve as a likely complement of the verb *sail*. Compared to control sentences with commas separating the clauses, sentences with plausible object incurred longer total reading times and more regressions on the main clause verb than did sentences with implausible objects (see also DeDe, 2010; Pickering, Traxler, & Crocker, 2000; Traxler, 2002). In sum, previous research bears out the effects of subcategorization, case marking, and plausibility on the reanalysis process of temporary object–subject ambiguities as in (1) in native processing.

L2 PROCESSING OF OBJECT–SUBJECT AND SUBJECT–OBJECT AMBIGUITIES

Several studies on L2 processing have examined how adult L2 speakers recruit subcategorization, morphosyntactic, and plausibility information in reading ambiguous sentences. In one of the earliest L2 sentence processing studies, Juffs and Harrington (1996) used a self-paced reading task to investigate how advanced Chinese learners of English and native controls process temporary object–subject ambiguities with obligatorily intransitive verbs (2a) and optionally transitive verbs (2b).

(2) a. After Sam arrived the guests began to eat and drink.
b. After Bill drank the water proved to be poisoned.

Like the natives, the L2 speakers showed reading delays on the main clause verb in sentences with optionally transitive verbs in the adjunct clause (2b) compared to (2a). These findings were replicated for a larger set of L1 Chinese, Japanese, and Spanish advanced learners of English in Juffs (2004), and studies on advanced English–French learners (Frenck-Mestre & Pynte, 1997), and English–Spanish (Jegerski, 2012) and Spanish–English readers (Dussias & Cramer-Scaltz, 2008) also found that L2 learners are sensitive to subcategorization information in the processing of object–subject ambiguities.

With respect to the L2 use of morphosyntax in subject–object ambiguities, Hopp (2006, 2010) reports that English and Dutch advanced learners of German did not use case information incrementally in disambiguating subject–object ambiguities in (3).

(3) Er denkt, dass der/den Physiker am Freitag den/der Chemiker gegrüßt hat. (SO/OS)
He thinks that the$_{NOM/ACC}$ physicist on Friday the$_{ACC/NOM}$ chemist greeted has
"He thinks the physicist greeted the chemist on Friday/the chemist greeted the physicist."

Only near-native speakers at the highest level of L2 end-state proficiency showed incremental effects of reanalysis in self-paced reading. In contrast, the advanced

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

learners treated the first noun of the embedded clause as being the subject, regardless of case information on the determiner. Comparable results were found for the disambiguation of object–subject ambiguities by subject–verb agreement (see also Havik et al., 2009) and for the processing of verbal inflection more generally in advanced L2 learners (e.g., Jiang, 2004, 2007; Rossi et al., 2006).[1]

For plausibility, Roberts and Felser (2011) tested object–subject garden path sentences in (4).

(4) a. The inspector warned the boss/the crimes would destroy many lives.
b. While the band played the song/the beer pleased all the customers.

Roberts and Felser studied L1 Greek advanced learners of English and native controls in self-paced reading. The L2 learners demonstrated effects of the plausibility of the NP only with the sentences in (4a). In (4b), the L2 learners showed local slowdowns on implausible object NPs, which indicates that they used plausibility information when incrementally integrating constituents in a parse. At the same time, there were no slowdowns for plausible versus implausible NPs on the region of the disambiguating verb or later. At first sight, this finding suggests that plausibility had no effect on recovery from garden paths. However, native speakers did not show significant reading time differences for either (4a) or (4b). Hence, failure to use plausibility information for reanalysis in this study did not seem to be specific to L2 processing and could have had a number of other causes. One option explored in a post hoc analysis by Roberts and Felser is that individual differences in reading speed might lead to variation in L2 processing. Overall reading speed interacted with garden path effects in both groups: Only the slower natives exhibited significant reading time differences according to plausibility, and the slower L2 readers were more committed to a plausible object analysis in that they started the reanalysis process later than faster readers. Roberts and Felser's data thus suggest that the extent to which L2 learners rely on plausibility information is subject to individual variation.

Individual differences in the L2 processing of ambiguous sentences have also been observed in other studies. Dussias and Pinar (2010) directly tested for the relation between working memory capacity and the use of plausibility information in *wh*-extraction sentences as in (5).

(5) a. Who did the police declare/know killed the pedestrian?
b. Who did the police declare/know the pedestrian killed?

Advanced Chinese–English learners had more difficulty processing subject extractions than object extractions in general. However, only the high-span L2 group demonstrated longer reading times after the verb in the plausible condition, indicating that they successfully recruited plausibility information for reanalyzing the sentence. The low-span group showed the inverse pattern, which indicates that they used plausibility information differently from their high-span counterparts or native readers (see also Williams, 2006). Taken together, the studies on the use of plausibility information in L2 processing hence suggest that L2 learners can readily employ plausibility to guide initial parsing commitments (see also

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Williams et al., 2001). At the same time, only L2 learners with greater resources, that is, high-span readers, or higher degrees of automaticity, that is faster readers, can use plausibility information successfully for revising an initial misanalysis.

Similar individual differences in reanalysis according to speed and working memory were reported for subcategorization and morphosyntax. For instance, Jegerski (2012) reported differences in reanalysis for sentences as in (2) according to reading speed in an L2 group of advanced English learners of Spanish, with the faster readers showing nativelike reading delays on the main clause verb with optionally transitive verbs compared to intransitive verbs. For morphosyntax, Jackson and Bobb (2009) tested advanced L1 English learners of German on subject–object ambiguities in *wh*-questions similar to the structures in (3). Although all L2 learners showed an initial subject-first preference in processing, there were no significant reading time differences on the regions where reanalysis becomes necessary. Moreover, memory span did not interact with reading times, with the exception that low-span speakers had greater difficulty recovering from an initial misanalysis at all.

Putting previous research in perspective, we conclude that L2 learners integrate subcategorization and plausibility information to inform their initial processing; however, the use of subcategorization and plausibility to guide reanalysis seems to be modulated by proficiency or reading speed and, with respect to *wh*-extraction, working memory differences. Finally, even advanced L2 learners do not reliably recruit morphosyntax, in this case inflectional morphology, for reanalysis. Although advances in proficiency are associated with increasing use of morphosyntax in incremental processing (Hopp, 2006, 2010), neither proficiency nor working memory seems to explain the range of interindividual variation.

Hence, previous research highlights that both specific resources like working memory and global markers of individual differences like proficiency or reading speed have effects on how L2 learners can revise an initial parsing commitment for reanalysis. It is as yet an open question to what extent these differences in resources affect L2 processing of different types of information in similar ways or whether the use of different types of information is mandated by different individual differences factors. This study intends to go some way toward addressing these issues by tackling the following questions.

- Which individual difference factors show effects for which type of information in L2 processing?
- Do these factors interact or is there a common cause for individual differences across information types?
- Do individual differences lead to qualitatively different processing patterns in the L2?

We investigate these issues with respect to temporary object–subject ambiguities as in (1). As discussed above, native and nonnative readers are garden-pathed with sentences such as *When the band played the song pleased everyone* in that they initially integrate a postverbal NP as the direct object of the verb and then need to revise this analysis. In the following study, we focus on which information types L2 learners recruit in revising the parse and the extent to which reanalysis is affected

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

by individual differences. We look at how reanalysis signaled by subcategorization, plausibility, and morphosyntactic information is affected by individual differences in proficiency, reading speed, working memory, automaticity of lexical access, and information integration ability.

THE PRESENT STUDY

To investigate the interaction of different information types in reanalysis from temporary object–subject ambiguities and individual differences in L2 processing, we created sets of garden path sentences as in (6).

(6) 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. *(Pronoun)*
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 (6a), no garden path arises, because punctuation prevents the integration of the NP *the boy* as the direct object of the verb *praying*. In all other sentences in (6), the absence of a comma allows for the provisional integration of the following constituent as the object of the adjunct clause verb, which subsequently needs to be revised on the main clause verb. In consequence, the focus of the present study is on the processing of the main clause verb, which provides unambiguous evidence for reanalysis.

As summarized above, the different information types biasing toward an object interpretation of the postverbal NP in (6b–6e) determine the strength of the reanalysis effect. In the intransitive condition (6b), subcategorization information on the verb signals that the following NP could not be its object. In the pronoun condition (6c), case marking on the pronoun indicates that the NP cannot be integrated as an object. In consequence, reanalysis costs on the main clause verb should be low in (6b and 6c). However, in the implausible (6d) and the plausible (6e) conditions, both the subcategorization properties of the verb and the morphosyntax of the NP license an object analysis of the postverbal NP; therefore, both sentences potentially invoke reanalysis. In (6d), plausibility information rules out the NP as a likely object of the verb. In contrast, the NP in (6e) constitutes a plausible object of the verb.

In addition, we test how individual differences affect the use of different information types for reanalysis in the sentences in (6) in L2 processing. If individual differences modulate reanalysis in temporary object–subject ambiguities, we expect to find interactions of reading times with individual scores on tests of proficiency, speed, cognitive capacity, automaticity, and integration ability.

First, proficiency constitutes a global measure of L2 mastery, and increases in proficiency are thus likely to go hand in hand with success in integrating different information types (e.g., Hoshino et al., 2010; McLaughlin et al., 2010). Second, similar to proficiency, reading speed reflects efficient computation at all levels of linguistic processing, such that it is expected to interact with the incremental use of any type of information. Third, working memory also handles the integration

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Table 1. *Participant characteristics*

	L2 Group ($n = 75$)			L1 Group ($n = 18$)		
	Mean	Range	*SD*	Mean	Range	*SD*
Age	22.7	18–29	2.59	23.2	19–33	4.2
Age of onset	10.6	10–14	0.84			
Length of exposure	12.2	6–19	2.72			
Proficiency (OPT grammar scores)	80.5	61–96	7.96			

Note: L2, second language; L1, first language; OPT, Oxford Placement Test.

of information in linguistic processing, so that lower span L2 learners might run out of resources in coordinating morphosyntactic (Havik et al., 2009; Jackson & Bobb, 2009; Keating, 2010) and plausibility information (Dussias & Pinar, 2010) in L2 processing. Fourth, Dekydtspotter, Schwartz, and Sprouse (2006) argue that learners whose basic linguistic processing skills in word recognition have been automatized (Segalowitz & Segalowitz, 1993) have more capacity for computing morphosyntactic information. Fourth, following suggestions that L2 learners may differ in their variable allocation of resources to either morphosyntactic or semantic–pragmatic aspects of the input (e.g., Hopp, 2009; Sorace, 2011), depending on how well readers can integrate semantic or syntactic information in context (Kilborn, 1992), we would expect a high score on semantic integration capacity to correlate with success in integrating lexical–semantic and plausibility information, whereas a high score on syntactic integration ability may relate to greater reliance on morphosyntax.

Participants

Seventy-eight German–English speakers, who all were students of English at a German university at the time of testing, were recruited to take part in the study for course credit. All participants had normal or corrected to normal vision. None was informed about the purpose of the study prior to the experiments. All participants reported having German as their first language (L1) and were first exposed to English at school. All had acquired English in an instructional setting, and only few participants stated having spent more than a few weeks in English-speaking countries (mean = 1.2 months). The participants hence formed a very homogeneous population in their linguistic, educational, and experiential backgrounds. One participant had to be excluded from analysis due to an incomplete data set in the eye tracking experiment; 2 participants were removed from further analysis because they did not complete all tests. The data of all remaining 75 participants were retained for analysis (see Table 1). To determine proficiency levels, the participants took the grammar part of the Oxford Placement Test (Allen, 1992), which is a test that has been widely used as an independent measure of proficiency in L2 processing studies (e.g., Felser, Roberts, Marinis, & Gross, 2003; Marinis et al., 2005; Roberts & Felser, 2011). Scores ranged from 61 to 96 out of

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

100, such that participants represented the categories from midintermediate to near-native speakers according to the Oxford Placement Test classification (Allen, 1992).

Eighteen native speakers of English, who were exchange students or beginning doctoral or postdoctoral students at a German university at the time of testing, constituted the native speaker control group. They were matched in mean age ($M = 23.2$) to the L2 group, and they were strongly dominant in English as per self-report. In a standardized 30-item proficiency test of German (Goethe Institute, 2010), they scored in the beginner to intermediate range ($M = 13$, $SD = 5.4$).

Materials and procedure

Eye tracking study. The experimental materials for the reading experiment encompassed 30 quintuplets of sentences as in (6) above (for all items, see Appendix A). The materials were adapted from a previous study on monolingual sentence processing using self-paced reading (Traxler, 2002). In the intransitive condition (6a and 6b), only obligatorily intransitive verbs were selected. To rule out effects of L1 subcategorization differences on the verb (e.g., Frenck-Mestre & Pynte, 1997; Dussias & Cramer-Scaltz, 2008), only verbs were chosen whose translation equivalents in German were also intransitive. For the conditions in (6c–6e), optionally transitive verbs were chosen based on the transitivity ratings reported in Traxler (2002). The translation equivalents of these verbs in German were also optionally transitive. For the plausibility manipulation, nouns were chosen that either could be plausible objects (6d) or would be implausible objects (6c). The implausible objects in (6c) were the same NPs as those used in the intransitive conditions (6a and 6b), such that the only difference between these conditions would be transitivity yet plausibility.

To ascertain whether the plausibility manipulation works for the bilingual population tested in this study, five German–English speakers drawn from the same population as those tested in the main study—but who did not take part in the main study—rated the plausibility of the sentences in (6d and 6d). The sentences were presented as single clauses (e.g., *The girl was playing the piano*) in a 60-item questionnaire on a 7-point scale, ranging from 1 (*very implausible*) to 7 (*very plausible*). The results show a statistically significant difference, $t\,(1, 29) = 26.940$, $p < .001$, between the sentences in (6d) and the sentences in (6e), with mean ratings of 1.69 and 6.42, respectively.

Finally, the case condition (6c) was constructed by replacing the postverbal noun by a third-person unambiguously nominative-marked pronoun (*he*, *she*, or *they*). In addition to the 30 experimental sentences, the experiment included 78 fillers. Twenty-four fillers were constructed similarly to the sentences in (6), yet they contained transitive verbs and a (plausible) full NP object in the preposed adjunct clause. We included these sentences to prevent participants from becoming aware of the experimental manipulation. The other fillers were of various grammatical structures, none of which contained intransitive verbs or object–subject ambiguities.

Comprehension questions followed all experimental items and 20 of the fillers. The comprehension questions corresponding to the experimental items always

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

targeted the subject of the main clause predicate (e.g., *Who made funny noises?*), presenting 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 check whether participants had successfully recovered from the garden path and correctly construed the NP as the subject of the main clause.

Five lists for the sentences in (6) were created in a Latin-Square design. All sentences were presented in random order, with a new randomization 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 eyetracker with a spatial resolution below 0.4 degrees. Tracking speed was 500 Hz. The sentences were presented in 20-point Arial in white on black on a 19-in. 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 11 and 20 min to complete the experiment.

For the reading time analysis, 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 analyzed first-pass reading time, second-pass reading time, total reading time, and number of regressions. First-pass reading time is the sum of all fixations in a region before the reader exits this region to the right for the first time. First-pass reading time has been taken to index first-pass processing (e.g., Rayner, 1998). (Nonzero) second-pass reading time refers to the summed length of all fixations in a region when the reader reenters this region from the right, provided he made a fixation. It can thus be interpreted as the effort associated with reanalysis when it is undertaken (e.g., Vasishth & Drenhaus, 2011). Total reading time is the sum total of all fixations in an area of interest. These later measures include rereading times of these regions and can thus be related to reanalysis and integration processes in text comprehension (e.g., Staub & Rayner, 2007). The number of regressions expresses how often readers launched a backward eye movement out of a particular region, a measure that can serve to indicate where reanalysis processes occur (e.g., Reichle, Rayner & Pollatsek, 2003). Finally, we calculated total reading speed as the sum of total reading time across all regions of interest and all conditions.

Working memory: Reading span task. All participants took the English-language reading span task developed by Ariji, Omaki, and Tatsuta (2003). In total, 70 sentences were presented in arrays of four sets of sentences in a computerized task using E-Prime 1.2 (Schneider, Eschman, & Zuccollotto, 2002). Set size increased from 2 to 5 sentences. Participants had to memorize one noun that was printed in capitals in each sentence and give a plausibility judgment after each sentence. Overall, half of the sentences were semantically plausible, the other half implausible.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Responses were coded as accurate if participants recalled the word in capitals and judged the sentence in which it had occurred correctly. The total reading span score was calculated as the sum of all accurate responses.[2]

Automaticity: Lexical decision task. In order to test the automaticity of basic lexical processing, a lexical decision task was designed following the rationale in Segalowitz and Segalowitz (1993; see also Segalowitz, Segalowitz, & Woods, 1998). Eighty words were chosen from the updated list of category norms in Van Overschelde, Rawson, and Dunlosky (2004). These words were taken from the two most frequently named English items in central semantic fields. Forty words, 1 from each semantic field, were assigned to the real-word condition. For the other 40 words, English pseudowords were designed by changing one syllable in accordance with English phonotactic constraints. In a lexical decision task implemented in E-Prime 1.2, all items were presented in random order in the center of the screen. Participants had to make a lexical decision within 5000 ms. Accuracy and response times were collected. Responses faster than 80 ms and slower than 1000 ms were discarded, which affected a total of 5.6% of the data for real words. As a measure of automaticity, the coefficient of variance was computed by dividing average reaction times to all real English words that had been judged correctly by the standard deviation of the reaction times for each participant, following the reasoning in Segalowitz and Segalowitz (1993). They argue that the coefficient of variance provides a superior measure of the automaticity of lexical processing than speed per se. Whereas decreases in reaction times may result from the speeding up of lexical recognition, no matter whether its processes are explicit or implicit, the coefficient of variance captures the degree of variance in lexical recognition relative to its overall speed. If variance is low, the process is likely to have been automatized (see also Harrington, 2007; Segalowitz & Hulstijn, 2005). The coefficient of variance was used as the measure of automaticity of lexical processing in all later analyses.

Information integration: Word monitoring task. To obtain independent measures of grammatical integration ability in context, a word monitoring task was administered based on the rationale in Kilborn (1992) and Oliver, Gullberg, Hellwig, Mitterer, and Indefrey (2012). Similar to the Kilborn study, the same 40 nouns from the category norms in Overschelde et al. (2004) were chosen. For each word, triplets of sentences were constructed (7). The target word is given in boldface in (7):

(7) a. Her boyfriend gave her a beautiful **diamond** for Christmas. (normal prose)
b. The gardener fly him the great **diamond** after mother. (syntactic prose)
c. Power it big rain a she **diamond** flower over. (random prose)

In (7a), both the semantic and the syntactic context are preserved, while in (7b) only the syntactic context remains because all words except for the target word have been replaced by words of the same lexical category. In (7c), neither semantic nor syntactic context are preserved. For half of the sentences, the target words were in the first half of the sentences, for the other half in the second part. The target

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

words in each triplet were in the same position of the string in all conditions. All sentences were spoken at a moderate pace by a native speaker of British English, who was naive as to the purpose of the experiment. The stimuli were digitally recorded and normalized in volume. Participants listened to the sentences over headphones.

In a word monitoring task implemented in E-Prime 1.2, participants saw the target word in the center of the screen and then pushed the space bar to start the auditory presentation of a sentence. They had to press the space bar again as soon as they heard the target word. After six practice items, the 40 sentences were presented in three blocks of normal (7a), syntactic (7b), and random (7c) prose; order within each block was fully randomized. Response times faster than 80 ms and longer than 1000 ms were removed from analysis; in total, this affected 7.4% of the data. Based on the average reaction times, a score of syntactic integration ability was computed by subtracting reaction times for each participant in the syntactic condition (7b) from those in the random condition (7c). The score was normalized by dividing it by the sum of the two reaction times. The syntactic integration score thus represents how well participants can exploit syntactic information in anticipating an upcoming target word. The higher the score, the greater the facilitation of additional syntactic information in anticipating words. Analogously, a score of semantic integration ability was computed by subtracting reaction times in the syntactic condition from those in the normal condition and normalizing them. The semantic integration score hence encodes how much speakers benefit from additional semantic information in predicting the target word. The normalized integration scores were used in all further analyses.

The participants were tested individually and took the reaction time experiments in a block of the following order: (a) lexical decision task, (b) reading span task, and (c) word monitoring task. About half of them completed the reading experiment and filled in questionnaires before the reaction time block, and the other half took the reading experiment after the other experiments. Each participant filled in the proficiency test on a separate occasion in class around the time of the main experiment.

Results for eye tracking (all participants). For all reading time measures, the data for a particular region were excluded if the reading time measure for that region was zero. In addition, fixations shorter than 80 ms were discarded, because readers cannot extract sufficient information in such short fixations (Rayner, 1998). In addition, all fixations longer than 2000 ms were excluded, because these most likely reflect tracker loss. In all, this affected less than 6% of all the data. As is common practice, only the trials for which the comprehension questions had been answered accurately were used for further analysis.

In the first step of the analysis, we present the results from the entire group of speakers to assess the processing pattern of the L2 group against the performance of the native control group. Table 2 gives accuracy in answering the comprehension questions and charts the response times to the comprehension questions that were responded to accurately.

The overall very high comprehension accuracy indicates that the participants read the sentences attentively and that they successfully reanalyzed any potential

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Table 2. *Accuracy and RTs (ms) of answers to comprehension questions for all L2 participants (n = 75) and all L1 participants (n = 18)*

	L2 Group ($n = 75$)		L1 Group ($n = 18$)	
	Accuracy (%)	RT (*SD*)	Accuracy (%)	RT (*SD*)
(8a) Comma	91.8	1966 (422)	92.6	1960 (1051)
(8b) Intransitive	94.2	2071 (591)	94.5	1941 (918)
(8c) Pronoun	94.4	2100 (502)	94.5	2232 (1027)
(8d) Implausible	94.4	2028 (475)	99.1	2149 (1111)
(8e) Plausible	89.6	2167 (662)	96.3	2257 (1155)

Note: RT, response time; L2, second language; L1, first language.

garden paths. There was no significant difference between the conditions in terms of response times in the L2 group, $F\ (4, 370) = 1.472$, $p = .210$, or the L1 group, $F\ (4, 89) = 2.026$, $p = .09$.

Table 3 gives the reading time measures for each condition and for each region of interest. For each region of interest, we report planned comparisons between the conditions based on previous research on monolingual speakers of English (e.g., Pickering & Traxler, 1998; Staub, 2007; Traxler & Pickering, 1996). The set of sentences in (6) is repeated here, as (8), for the reader's convenience.

(8) 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. *(Pronoun)*
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)*

First, to establish whether readers are sensitive to punctuation in the control condition, we compare (8a) and (8b), henceforth termed the *control comparison*. Reading times are expected to be longer in (8b) on the NP for lack of orthographic evidence that a new clause needs to be initiated (e.g., Staub, 2007).

Second, to identify effects of transitivity (*transitivity comparison*), comparisons are run between (8b) and (8d). If readers recruit transitivity information immediately, we would expect differences in first-pass parsing on the NP between (8b) and (8d). If subcategorization properties of the verb affect reanalysis, reading times on the main clause verb should differ, with shorter reading times and fewer regressions in the intransitive (8b) than the transitive (8d) condition (e.g., Staub, 2007; van Gompel & Pickering, 2001).

Third, to see whether readers make a difference according to case (*case comparison*), we compare (8c) and (8a).[3] Because reading times on the pronoun are likely to be shorter than on full NPs due to the differences in length, we cannot make direct comparisons on the NP region. Hence, we focus on the region of the main clause verb. If case is used immediately for reanalyzing subject–object

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Table 3. *Reading times (ms) and number of regressions*

		Adjunct Clause Verb		Noun Phrase		Main Clause Verb		Final Region	
		L2	L1	L2	L1	L2	L1	L2	L1
(8a) Comma	First pass	289 (9)	237 (13)	309 (8)	242 (12)	277 (7)	249 (14)	663 (19)	450 (28)
	Second pass	274 (15)	309 (26)	324 (18)	290 (26)	306 (23)	345 (42)	455 (32)	361 (38)
	Total	390 (12)	347 (22)	416 (13)	398 (25)	372 (13)	415 (31)	783 (20)	574 (31)
	Regressions	102	42	58	21	76	36	281	101
(8b) Intransitive	First pass	267 (7)	239 (18)	379 (10)	295 (16)	281 (8)	237 (13)	652 (19)	477 (37)
	Second pass	364 (22)	421 (37)	365 (21)	401 (33)	297 (16)	331 (30)	520 (34)	402 (34)
	Total	401 (16)	517 (37)	512 (15)	521 (33)	392 (12)	391 (25)	783 (23)	621 (39)
	Regressions	75	53	103	58	62	37	256	93
(8c) Pronoun	First pass	254 (6)	257 (16)	256 (7)	223 (12)	297 (9)	241 (13)	669 (21)	459 (36)
	Second pass	359 (22)	344 (31)	299 (23)	275 (29)	342 (20)	309 (28)	530 (31)	438 (51)
	Total	419 (15)	488 (32)	363 (14)	355 (23)	429 (14)	410 (26)	864 (28)	694 (50)
	Regressions	88	55	54	42	116	51	305	113
(8d) Implausible	First pass	257 (7)	239 (12)	371 (9)	313 (16)	298 (8)	263 (13)	669 (20)	482 (34)
	Second pass	339 (16)	512 (47)	431 (22)	410 (35)	322 (17)	340 (31)	466 (31)	388 (39)
	Total	415 (15)	562 (41)	569 (18)	565 (35)	428 (13)	447 (27)	817 (23)	638 (41)
	Regressions	88	84	101	53	100	42	292	107
(8e) Plausible	First pass	258 (71)	250 (14)	369 (10)	277 (12)	309 (8)	292 (15)	620 (20)	411 (30)
	Second pass	354 (17)	500 (45)	445 (24)	533 (39)	386 (19)	483 (38)	604 (33)	537 (53)
	Total	423 (14)	627 (51)	597 (20)	669 (40)	490 (16)	580 (38)	868 (26)	735 (54)
	Regressions	105	68	116	79	125	68	300	128

Note: The regions of interest are delimited by brackets: When the girl [was playing] [the boy] [made] [some funny noises]. Standard errors are given in parentheses. All participants, L2: $n = 75$; L1: $n = 18$. L2, second language; L1, first language.

ambiguities, we would expect there to be no difference between the control condition (8a) and the pronoun condition (8c) for reading times (Traxler & Pickering, 1996).

Fourth, to identify effects of plausibility (*plausibility comparison*), (8d) and (8e) are compared. If plausibility information guides initial parsing, reading times on the NP should be shorter in the plausible condition (8e) than in the implausible condition (8d), because a plausible NP can be integrated as an object more easily. For reanalysis, however, the inverse pattern is expected. Implausible NPs should facilitate recovery from the erroneous object analysis, because the commitment to a direct object analysis is weaker with implausible than plausible NPs (e.g., Pickering & Traxler, 1998).

Given the drastically different group sizes, the data from the L2 learners and the natives were analyzed in two separate linear mixed-effect models (see Baayen, Davidson, & Bates, 2008). Unlike traditional analyses of variance (ANOVAs), mixed-effects models use unaveraged data as input and incorporate random effects of both participants and items within a single analysis. The present analyses included subject and item intercepts and slopes as random effects and condition as a fixed effect. In the following, we present the results of all analyses for which the effect of condition reached significance for each region of interest in turn. Table 4 charts all statistical comparisons, reporting regression coefficients, standard errors, Wald z values, and the corresponding p values.

Control comparison. For the control comparison, both the natives and the L2 group demonstrate significant differences between the comma condition (8a) and the intransitive condition (8b) in the adjunct verb region and the NP region across several measures (Table 4). All of these differences point to greater rereading of the first verb and the NP in the ambiguous no-comma condition bearing out the effects of punctuation, which signals the beginning of a new clause.

Transitivity comparison. On the NP, the L2 group have longer second-pass (365 vs. 431 ms) and total reading times (512 vs. 569 ms) for optionally transitive verbs (8d) compared to obligatorily intransitive verbs (8b). In the subsequent main clause verb region, there are significant differences in total reading times (392 vs. 428 ms) and the number of regressions (62 vs. 100), which bear out the effects of transitivity. Although the natives also show numerically longer total reading times for optionally transitive than for intransitive verbs in the NP region (521 vs. 565 ms), only the difference in the main clause verb region (391 vs. 447 ms) reaches statistical significance.

Case comparison. On the critical main clause verb region, the natives do not show any significant differences in reading times or regressions between the comma condition (8a) and the nominative pronoun condition (8c). In the final region, (8a) and (8c) differ only in total reading times (574 vs. 694 ms). This difference may be taken to reflect longer wrap-up effects for integrating the interpretation of sentences including a pronoun with a sentence-external referent in (8c) versus (8a), which does not contain any pronouns.

In contrast, the L2 group have significantly longer total reading times (372 vs. 429 ms) and more regressions (76 vs. 116) on the main clause verb for the pronoun

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Table 4. *Effects of condition for all eye movement measures in four regions of interest*

		Adjunct Clause Verb		Noun Phrase		Main Clause Verb		Final Region	
		L2	L1	L2	L1	L2	L1	L2	L1
(8a) & (8b) Control	First pass	21.4/10.3; 2.087*		68.7/11.6; 5.940***	54.8/17.1; 3.214**				
	Second pass	95.2/24.1; 3.947***	116.3/49.3; 2.358*	38.4/26.3; 1.457	112.9/42.6; 2.650**			66.6/45.3; 1.470	
	Total		167.3/35.8; 4.667***	95.1/16.5; 5.775***	129.8/37.9; 3.423**				
	Regressions	0.06/0.03; 2.087*		0.11/0.03; 3.580**	0.36/0.09; 4.098***	0.04/0.03; 1.239	<1	0.07/0.04; 1.515	
(8b) & (8d) Transitivity	First pass					17.4/9.7; 1.802	26.6/16.9; 1.574		
	Second pass	27.7/23.7; 1.169	97.0/54.7; 1.774	67.4/28.0; 2.404*		30.6/22.9; 1.332		62.6/45.0; 1.391	
	Total		56.5/44.7; 1.264	64.3/19.9; 3.238**	50.8/41.2; 1.235	59.7/15.4; 3.882***	60.1/30.8; 1.951*	32.2/22.8; 1.412	
	Regressions		0.28/0.11; 2.595*			0.09/0.03; 2.948**		0.08/0.42; 1.872	
(8a) & (8c) Case	First pass	35.3/10.4; 3.396***		55.5/10.2; 5.454***	18.8/16.2; 1.160	19.2/10.7; 1.801			
	Second pass	88.2/27.4; 3.222***				43.7/29.3; 1.491		80.0/43.7; 1.831	77.9/64.8; 1.148
	Total	28.3/17.2; 1.645	142.2/35.4; 4.019***	57.5/16.7; 3.441***	46.1/32.9; 1.402	56.8/17.1; 3.333***		84.8/25.5; 3.329**	122.8/50.7; 2.421*
	Regressions		0.12/0.09; 1.236		0.20/0.08; 2.632*	0.10/0.03; 2.860**	0.14/0.09; 1.619	0.06/0.05; 1.358	
(8d) & (8e) Plausibility	First pass				36.6/19.5; 1.879	11.5/10.6; 1.082		46.3/22.5; 2.013*	61.6/36.3; 1.699
	Second pass				119.4/51.2; 2.330*	65.6/25.0; 2.621**	140.3/49.2; 2.851**	140.8/41.1; 3.421***	137.3/65.0; 2.115*
	Total		54.4/53.9; 1.009	2.235	95.5/44.2; 2.159*	63.8/18.3; 3.488***	124.1/41.8; 2.969**	59.5/26.4; 2.249*	160.0/52.1; 2.033*
	Regressions	0.04/0.03; 1.403	0.15/0.11; 1.372	1.406	0.24/0.09; 2.522*	0.06/0.03; 1.967*	0.23/0.09; 2.465*	0.17/0.05; 3.570***	0.22/0.12; 1.893

Note: The regions of interest are delimited by brackets: When the girl [was playing] [the boy] [made] [some funny noises]. All participants, L2 group: $n = 75$, and L1 group: $n = 18$. The values are coefficient estimate β/standard error $SE(\beta)$ with the associated Wald z score. L2, second language; L1, first language.
*$p < .05$. **$p < .01$. ***$p < .001$.

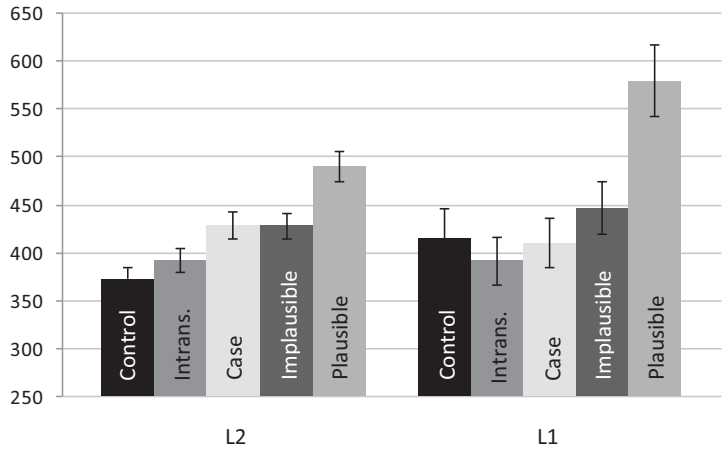


Figure 1. The total reading times (ms) in the main clause verb region (including standard error) for the second language (L2) group ($n = 75$) and first language (L1) group ($n = 18$).

sentence (8c) than for the comma condition (8a). In the final region, total reading times (783 vs. 864 ms) are also longer for (8c) than (8a), as in the native control group. The longer reading times and higher number of regressions for the pronoun sentences compared to the control condition on the main clause verb indicate that reanalysis from an initial object interpretation is undertaken by the L2 group despite nominative case marking on the pronoun.

Plausibility comparison. In the main clause verb region, both the L1 and the L2 groups make significant differences between the implausible object condition (8d) and the plausible object condition (8e) in second-pass reading times (L1: 340 vs. 482 ms; L2: 322 vs. 386 ms), total reading times (L1: 447 vs. 580 ms; L2: 428 vs. 490 ms), and in the number of regressions (L1: 42 vs. 68; L2: 100 vs. 125). In the final region, differences are significant in second-pass reading times (L1: 388 vs. 537 ms; L2: 466 vs. 604 ms) and total reading times (L1: 638 vs. 735 ms; L2: 817 vs. 868 ms). Across groups, plausible NPs have longer reading times and more regressions than implausible NPs as objects.

To illustrate the processing patterns, Figure 1 graphs the total reading times on the main clause verb for the L2 group and for the L1 group.

Discussion

The results for the entire group of L1 and L2 readers can be summarized as follows. For the control comparison, native and nonnative readers take more time on the NP in first-pass processing as shown by the significant differences between the comma condition (8a) and the intransitive condition (8b). Similar to previous eyetracking studies on native speakers (Pickering & Traxler, 1998, Experiments 1 & 3; Staub, 2007), this delay illustrates that native and nonnative readers take additional time

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

to open a new clause in the absence of punctuation that provides information on clausal segmentation. For the purpose of the following comparisons, this finding underlines that the control sentence in (8a) constitutes a neutral baseline condition that does not involve reanalysis.

For the no-comma sentences in (8b–8e), there were no differences in first-pass reading times on the NP depending on verb transitivity or plausibility in either the L1 or the L2 group in the NP region, with the exception of the orthographically shorter pronouns in (8c). This null effect in both groups differs from findings in previous research on monolinguals. For subcategorization, van Gompel and Pickering (2001) and Staub (2007) found longer first-pass reading measures on NPs following intransitive than optionally transitive verbs. For plausibility, Pickering and Traxler (1998) report that readers take significantly less time to read a plausible NP in postverbal position than an implausible NP. The fact that this study did not unearth effects in first-pass reading times on the postverbal NP is likely due to differences in the sentence stimuli compared to previous studies on monolingual processing. All of these studies used complex NPs containing prepositional phrases (e.g., *the magazine about fishing*), which extended the temporary ambiguity across the prepositional phrases. By contrast, the present study employed only simple NPs or pronouns. The lack of differences between the conditions on the NP region suggests that readers may have previewed the disambiguating region of the main clause verb or proceeded to it before they recruited subcategorization or plausibility information in first-pass processing. In this respect, the present data are similar to the findings in Roberts and Felser (2011), who also used simple NPs in a self-paced reading study on sentences like (8b–8d) and did not find garden path effects in natives and a group of (faster) L2 readers on the NP region.

In the subsequent main clause regions, few significant effects could be found in the first-pass data, as in previous eyetracking studies with natives (Pickering & Traxler, 1998). Instead, the effects of the experimental manipulations in the no-comma conditions are mainly visible in later measures associated with reanalysis, in particular, total reading times and in the number of regressions. Here, the L2 group show marked differences from the native control group in how it recruits subcategorization, case marking, and plausibility information.

Whereas the natives demonstrate reading time differences according to subcategorization, case marking, as well as plausibility, the L2 group only show sensitivity to subcategorization and plausibility information in on-line reading.

For *transitivity*, the L2 group, like the native controls, exhibit significant differences in total reading times on the main clause verb, with reading times being shorter for sentences containing intransitive verbs (8b) than transitive verbs (8d). For *plausibility*, the significant differences between the implausible (8d) and the plausible (8e) object conditions on the main clause verb and the final region bear out that nonnative readers, like natives, recruit plausibility information in L2 processing because sentences with plausible NPs elicit longer reading times than do sentences with implausible NPs.

For *case*, the native control group do not show any significant differences between the pronoun condition (8c) and the comma condition (8a) on the main clause region, which indicates that case marking information is recruited immediately for revising the parse (see also Traxler & Pickering, 1996). In contrast, for

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

the L2 group, the significant differences on the main clause verb region suggest that L2 readers do not readily integrate case information because the pronoun sentences (8c) require some reanalysis compared to the unambiguous sentences (8a). Further, for the L2 group, the total reading times in the pronoun condition (8c, 429 ms) and the implausible object condition (8d, 428 ms) as well as the number of regressions (116 vs. 100) are statistically indistinguishable (all zs < 1). Because the processing patterns did not differ between the pronoun (8c) and the implausible object condition (8d), case, although saliently marked on pronouns, can neither be incrementally used for avoiding initial misanalysis nor does it aid in revising the initial object analysis of the postverbal NP.

In sum, the processing patterns attested in the present experiment for the entire L2 group replicate and add to previous findings on L2 sentence processing. In line with other studies, we find that advanced L2 learners make reliable use of subcategorization information (see also Dussias & Cramer-Scaltz, 2008; Frenck-Mestre & Pynte, 1997) and plausibility information (see also Dussias & Pinar, 2010; Roberts & Felser, 2011). At the same time, L2 learners have difficulty recruiting morphosyntactic information incrementally in on-line comprehension. This finding resonates with previous research on advanced L2 learners (e.g., Hopp, 2006; Jackson & Dussias, 2009) that did not find incremental sensitivity to case marking for reanalysis in non-near-native L2 learners. It also supplements studies comparing different aspects of L2 processing in that morphosyntax, yet not semantic or pragmatic information, poses persistent difficulties in L2 processing (e.g., Felser et al., 2003; Hopp, 2007). Having established general nonnative versus native differences at the group level, we probe the extent to which the L2 processing patterns are moderated by individual differences in the next section.

INDIVIDUAL DIFFERENCES

In this section, we first present the results from the respective experiments on individual differences; second, we explore their interrelations; and third, we relate them to the reading time data.[4] Table 5 shows the findings from the tasks probing individual differences in L2 processing.

Although all 75 participants came from a similar population and had intermediate to near-native proficiency in English, cognitive as well as linguistic variables exhibit a considerable range of variation. To gauge the extent to which these variables are related, Table 6 presents bivariate correlations between the individual differences variables.

Table 6 illustrates that working memory and proficiency score are moderately correlated (see also Jackson & Bobb, 2009; Miyake & Friedman, 1998). Because working memory capacity contributes to language proficiency, yet proficiency encompasses more general mastery of the L2, both variables were retained for further analyses. Further, the scores for semantic and syntactic integration ability exhibit a weak negative correlation. Such a correlation indexes that L2 learners who demonstrate greater gains for word recognition in context in exploiting syntactic information show smaller gains in exploiting semantic information and vice versa. In this respect, it is interesting to note that the scores for automaticity in lexical access and for sentence integration ability do not correlate. Even though both

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Table 5. *Results from individual differences tasks*

		L2 ($n = 75$)			L1 ($n = 18$)		
	Type	Mean	*SD*	Range	Mean	*SD*	Range
Proficiency	Score (max. 100)	80.5	7.96	61–96			
Working memory	Score (max. 70)	33.7	8.38	17–56	47	9.2	33–63
Lexical access	Coefficient of variation	0.246	0.076	0.117–0.407	0.171	0.130	0.130–0.216
Semantic integration	Normalized score	–0.058	0.058	−0.169–0.140	–0.134	0.069	−0.315–0.006
Syntactic integration	Normalized score	0.061	0.047	−0.052–0.171	0.086	0.062	−0.005–0.208
Reading speed	Overall reading time (ms)	500	164	208–1230	489	143	237–761

Note: All L2 participants ($n = 75$) and all natives ($n = 18$). L2, second language; L1, first language.

Table 6. *Bivariate correlations between individual differences variables*

	Working Memory	Lexical Access	Semantic Integration	Syntactic Integration	Reading Speed
Proficiency	.444**	−.192	.089	.130	−.280*
Working memory		−.126	.110	−.006	−.272*
Lexical access			−.203	−.027	.250*
Semantic integration				−.289*	−.270*
Syntactic integration					−.028

Note: All are second language participants ($n = 75$).
*$p < .05$. **$p < .01$.

tasks require the identification of the same words, the scores for word recognition outside of context and for word monitoring within a sentence context are different. In other words, they likely refer to different underlying processes. Finally, reading speed correlates moderately with all other variables, except for syntactic integration. These correlations underline the global nature of the speed measure, which encompasses quick and efficient computation at all levels of linguistic analysis. Because none of the weak associations between individual difference variables presents a threat to a regression analysis in terms of (multi)collinearity, each of the six individual differences variables was retained as a predictor variable for the regression analysis reported in the following section.

Individual differences and reading comprehension

In this section, we report whether the processing of temporary object–subject ambiguities in (8) is systematically affected by individual differences. In order to search for interactions with individual differences, we analyzed the respective influence of proficiency, reading speed, working memory, lexical automaticity, and syntactic as well as semantic integration ability on the effects of case, transitivity, and plausibility discovered in the analysis of the entire group. We focus on the region of the main clause verb region, because this is where significant effects came out in the group analysis. Linear mixed-effects regression analyses were conducted with participants and items as random variables and condition (baseline vs. intransitive, pronoun, implausible, and plausible, respectively) and the individual difference factors (proficiency, reading span, total speed, coefficient of variance in lexical decision, syntactic integration score, and semantic integration score) as fixed-effect factors for reading times (first-pass reading times, second-pass reading times, and total reading times) and the number of regressions. Each potentially ambiguous condition in (8b–8e) was compared to the unambiguous comma condition in (8a) so that the analysis would encompass all effects of processing difficulty in the respective ambiguous conditions relative to an unambiguous control condition.

Initially, a model that only included the random factors (participants and items) was applied. This initial model was next enriched by adding the first fixed factor (i.e., Condition) and subsequently by including the individual differences

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

predictors as fixed factors one by one. Finally, the interactions of individual difference variables with the factor Condition were added to the model in stepwise fashion. Each successive pair of models was evaluated to assess whether the additional factor improved the model fit to the data, based on a series of log-likelihood ratio tests. The most complex model that significantly improved the fit over the previous model was considered to be the best fitting model, and its estimates are reported below. Table 7 reports all significant interactions of the individual differences factors with the factor Condition in the comparisons for transitivity, pronoun, implausible objects, and plausible objects.

Table 7 reveals three distinct patterns of interactions across the potentially ambiguous sentence types with respect to the unambiguous control condition. First, there are significant associations in first-pass reading time with the coefficient of variance in the lexical decision task. The interaction reaches significance for the implausible and (marginally) for the intransitive condition. The interactions each reflect negative correlations between the coefficient of variance and the difference between the experimental and the control conditions. Hence, L2 speakers with higher degrees of automaticity in lexical access show more pronounced effects of implausibility and subcategorization sooner in first-pass measures than do L2 learners with less routinized lexical access. However, differences in the automaticity of lexical access do not continue to affect sentence processing in later measures of processing associated with sentence integration and reanalysis.

Second, overall reading speed shows marginally significant interactions with second-pass and total reading times as well as the number of regressions in the plausible condition. The present study thus replicates the findings in Roberts and Felser (2011), who report effects of speed on the reanalysis of object–subject ambiguities that differ in plausibility. In both studies, overall speed constitutes an efficiency measure of how quickly readers can recover from a strong garden path effect by undoing a commitment to an initially plausible analysis.

Third, interactions with the syntactic integration score reach significance in first-pass reading times for the implausible condition and in second-pass reading times for the intransitive, the implausible, and (marginally) for the pronoun condition. Of all the individual difference variables, syntactic integration thus shows the most interactions in the reading time data. These associations hold for second-pass measures that are typically associated with reanalysis and integration processes (e.g., Reichle et al., 2003). Because the interactions suggest that syntactic integration ability constitutes an individual difference variable that modulates recovery processes in the L2 comprehension of object–subject ambiguities, we analyze the impact of syntactic integration in more detail.

In order to establish the nature of the interactions with syntactic integration, the L2 group was split into three groups of 25 participants each by their syntactic integration scores. We opted for a three-way division because the standard way of splitting the group by its median in previous L2 processing studies typically results in rather homogeneous groups and in misclassifications at the margins, so that effects of the dividing variable are often not observed (for discussion, see, e.g., Conway et al., 2005). Given that the present study affords enough participants to allow for a three-way division, we divided the participants into low, mid, and high syntactic integration subgroups (see also Long & Prat, 2008; Swets, Desmet,

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Table 7. *Interactions of reading measures with individual differences on main clause verb region*

Main Verb Region (8a) Versus	First-Pass Reading Time	Second-Pass Reading Time	Total Reading Times	No. of Regressions
(8b) Intransitive	$LDT_{CV} = 223.2/122$; $z = 1.830$†	$Int_{syn} = 1169.3/625$; $z = 1.872$*		WM = 0.06/0.04; $z = 1.701$†
(8c) Pronoun		$Int_{syn} = 471.5/280.4$; $z = 1.681$†		
(8d) Implausible	$LDT_{CV} = 289.1/130.1$; $z = 2.222$*	$Int_{syn} = 1252.3/602.9$; $z = 2.077$*		
	$Int_{syn} = 501.3/208.2$; $z = 2.407$*			
(8e) Plausible		Speed = 0.27/0.16; $z = 1.736$†	Speed = 0.2/0.12; $z = 1.778$†	Speed = 0.004/0.002; $z = 1.693$†

Note: All are second language participants ($n = 75$). The values are coefficient estimate β/standard error $SE(\beta)$ with the associated Wald z score. LDT_{CV}, lexical decision task/coefficient of variance; Int_{syn}, syntactic integration score; WM, working memory.
†$p < .1$. *$p < .05$.

Table 8. *Accuracy in answers to comprehension questions*

	Low Integration		Mid Integration		High Integration	
	Accuracy (%)	RT (ms)	Accuracy (%)	RT (ms)	Accuracy (%)	RT (ms)
(8a) Comma	92.7	1982	95.3	1990	91.4	1926
(8b) Intransitive	98.0	2134	84.7	2115	94.1	2091
(8c) Pronoun	95.2	2150	91.4	2063	90.9	1976
(8d) Implausible	96.6	2063	93.5	2065	90.2	1958
(8e) Plausible	94.6	2396	87.5	2048	94.7	2065

Note: Each second language group ($n = 25$) by syntactic integration score. RT, response time.

Hambrick, & Ferreira, 2007). The reading times for the respective comparisons were then computed individually for each group. In order to capture the full processing patterns of each group, all reading time measures are reported in the analyses by group, even though syntactic integration ability interacted significantly only with some of the measures in the global comparison. Before presenting the reading time comparisons, the groups were compared with respect to the individual differences variables collected in this study. One-way ANOVAs on the individual difference variables with group as a between-subjects factor do not yield significant effects for any measure, except for syntactic integration ability, $F(2, 72) = 128.327$, $p < .001$, and semantic integration ability, $F(2, 72) = 3.470$, $p = .036$, due to the intercorrelation of these factors (see Table 6). On all other measures, the groups are comparable in their characteristics, such that the three-way division according to integration scores is not confounded with effects of other individual difference variables.

Comprehension accuracy and response times

Table 8 gives accuracy rates and response times of the answers to the comprehension questions. Overall, comprehension accuracy is very high in all groups. The mid integration group has slightly depressed comprehension accuracy scores for sentences with intransitive verbs (8b) and for sentences with plausible NPs (8e). Because the group has the shortest overall reading times of all the groups, the slightly depressed accuracy rate might represent a speed–accuracy trade-off. In the response times, there are no significant effects of condition in any group (all Fs < 1).

Reading time analysis

For all three groups, Table 9 charts all statistical comparisons on first-pass, second-pass, and total reading times as well as the number of regressions by giving

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Table 9. *Effects of condition for all eye movement measures in three regions of interest by L2 syntactic integration group (n = 25 each)*

		Low Integration			Mid Integration			High Integration		
		Noun Phrase	Main Clause Verb	Final Region	Noun Phrase	Main Clause Verb	Final Region	Noun Phrase	Main Clause Verb	Final Region
(8b) & (8d) Transitivity	First pass		18.0/16.5; 1.089	49.5/38.1; 1.298		23.9/17.9; 1.336	54.7/41.9; 1.304	25.3/23.3; 1.085	42.6/16.4; 2.600*	61.7/33.4; 1.846†
	Second pass	96.6/52.6; 1.836†			99.4/55.0; 1.805†		101.6/94.1; 1.079			117.4/62.3; 1.885†
	Total	84.3/31.9; 2.642**			75.9/34.4; 2.203*	71.8/27.2; 2.636*	82.0/41.9; 1.958†		78.1/24.2; 3.227 **	53.4/38.9; 1.376
	Regressions		0.07/0.04; 1.673†	0.09/0.07; 1.338		0.08/0.06; 1.365	0.12/0.07; 1.775†		0.11/0.05; 2.061	
(8a) & (8c) Case	First pass	74.0/17.6; 4.209***			10.151***	40.7/18.2; 2.238*		33.3/16.9; 1.970*		
	Second pass		75.3/47.1; 1.600			74.1/54.1; 1.368		64.3/38.8; 1.656		98.7/78.8; 1.253
	Total		74.2/29.5; 2.511*	80.1/39.6; 2.021*	49.7/26.8; 1.858†	76.2/29.9; 2.547*	106.1/44.3; 2.392*	66.0/26.5; 2.490*		96.9/1.6; 1.880†
	Regressions		0.14/0.06; 2.441*			0.12/0.06; 1.951†				
(8d) & (8e) Plausibility	First pass		38.9/17.7; 2.205*							49.6/37.9; 1.309
	Second pass		85.0/40.7; 2.086*	116.0/68.9; 1.685†			119.6/81.1; 1.474	71.0/45.5; 1.560	104.4/38.4; 2.721*	150.1/67.6; 2.221*
	Total	90.7/37.9; 2.391*	117.8/32.0; 3.676***	60.8/43.9; 1.385					70.5/31.0; 2.277*	86.7/47.4; 1.830†
	Regressions	0.13/0.06; 2.244*	0.14/0.06; 2.476*	0.20/0.08; 2.467*			0.12/0.08; 1.565			0.18/0.08; 2.222*

Note: The regions of interest are delimited by brackets: When the girl was playing [the boy] [made] [some funny noises]. Standard errors are given in parentheses. The values are coefficient estimate β/standard error *SE*(β) with the associated Wald *z* score.
†$p < .1$. *$p < .05$. **$p < .01$. ***$p < .001$.

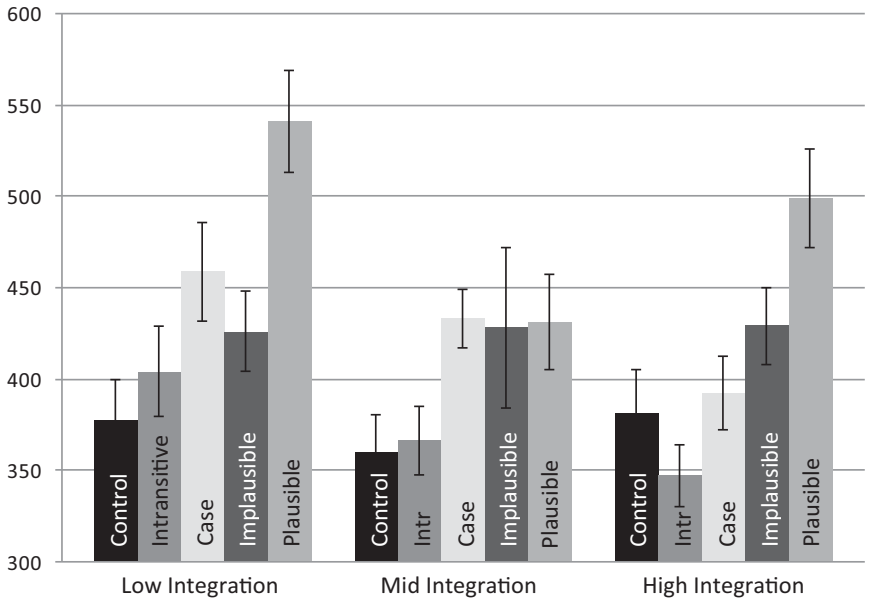


Figure 2. The total reading times (ms) in the main clause verb region (including standard error) for the low, mid, and high syntactic integration group ($n = 25$ each).

regression coefficients, standard errors, Wald z values, and the corresponding p values.

Figure 2 graphs the total reading times in the main clause verb region for each group. In the following, we report the reading time analysis for each group. The full data set of reading times and standard errors for each of the three groups is listed in Appendix A. In the following, we discuss the processing patterns for each group for the transitivity, the case, and the plausibility comparisons, respectively.

Low integration group. For the *transitivity comparison*, total reading times differ significantly (404 vs. 429 ms) for intransitive (8b) compared to optionally transitive verbs (8d) on the NP. There are no significant differences on any measure between (8b) and (8d) on the main clause verb or the final region.

For the *case comparison*, the low integration group does not show differences on the main clause verb in first-pass reading times between the comma condition (8a) and the pronoun condition (8c), yet it exhibits significantly longer total reading times (377 vs. 459 ms) and a higher number of regressions (19 vs. 39) in the pronoun condition. In the final region, total reading times also differ (827 vs. 908 ms).

For the *plausibility comparison*, the low integration group shows significant longer reading of the main clause verb in the implausible (8d) than in the plausible object condition (8e) in first-pass (287 vs. 327 ms) and second-pass reading

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

times (342 vs. 415 ms). Significant differences also obtain in total reading times (426 vs. 541 ms) and the number of regressions in the main clause verb region (28 vs. 46) and in the final region (102 vs. 128).

Mid integration group. For the *transitivity comparison* between the intransitive and the implausible object conditions (8b and 8d), the midintegration group shows a significant difference on the NP region in total reading times (510 vs. 570 ms). On the main clause verb, total reading times show a significant difference (366 vs. 428 ms) with sentences in the intransitive condition taking less time to read than those in the transitive condition. There are no significant effects in the final region.

For the *case comparison*, the mid integration group makes a significant difference on the main clause verb between the comma condition (8a) and the pronoun condition (8c) in first-pass reading times (270 vs. 311 ms), yet not in second-pass reading times (341 vs. 345 ms) and only marginally in the number of regressions (31 vs. 46). Total reading times also show a significant difference (360 vs. 428 ms). In the final region, there are only effects in total reading times, with the pronoun sentences taking longer to read than the control sentences (748 vs. 817 ms).

For the *plausibility comparison*, the mid integration group does not exhibit significant differences on any measure in any region.

High integration group. For the *transitivity comparison*, there are no significant differences between the intransitive (8b) and the implausible object condition (8d) on any measure in the NP region. In the main clause verb region, the high integration group exhibits significantly longer first-pass reading times (263 vs. 311 ms) and total reading times (347 vs. 429 ms) as well as more regressions (21 vs. 37) for sentences with optionally transitive verbs (8d) than for sentences with intransitive verbs (8b).

For the *case comparison*, there are no significant differences between the comma condition (8a) and the pronoun condition (8c) in any measures on the main clause verb. In the final region, only total reading times are marginally elevated in (8c) compared to (8a).

For the *plausibility comparison*, the high integration group makes no difference between the implausible (8d) and the plausible verb condition (8e) on the main clause verb in first-pass reading times (311 vs. 301 ms). Second-pass reading times are significantly faster in the implausible condition (296 vs. 400 ms) as are total reading times (429 vs. 499 ms). These effects spill over into the final region.

In summary, the three groups show distinct patterns in reanalyzing object–subject ambiguities—in particular with respect to the use of case marking and the use of plausibility information.

In order to establish whether there are interactions of case and plausibility with syntactic integration score, we further divided the sentences in the pronoun condition (8c) into two categories: sentences in which the pronouns could be plausible objects of the adjunct clause verb (e.g., *When Mary drove he*) and sentences for which the pronouns would be implausible objects (e.g., *When Mary baked he*). Although morphosyntactically inaccessible as objects, the pronouns would constitute plausible objects of the adjunct clause verb for 15 sentences; for

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

the other 15, they would be implausible objects (see Appendix A). Ten German–English speakers from a similar population to the one tested in the main study rated the plausibility of the pronoun as an object of the adjunct clause verb. In an off-line questionnaire, all 30 sentences were administered as single clauses with accusative object pronouns (e.g., *Mary baked him*). On a 7-point scale, raters made a significant difference, t (1, 29) = 18.462, $p < .001$, between sentences categorized as having plausible pronoun objects (mean rating = 5.5) and sentences with implausible pronoun objects (mean rating = 1.9). The post hoc division of the pronoun sentences in (8c) thus leads to groups of plausible pronoun objects and implausible pronoun, respectively.

To explore the interactions of case and plausibility, we computed a mixed model ANOVA on the total reading times on the main clause verb region with plausibility (plausible vs. implausible), NP type (pronoun vs. full NP), and syntactic integration score as fixed factors and subjects and items as random factors. The analysis yields a significant main effect of plausibility ($\beta = 104.4$, $SE = 51.1$, $z = 2.042$; $p = .042$), yet not of NP type ($\beta = 119.1$, $SE = 85.0$, $z = 1.401$; $p = .161$) nor an interaction of plausibility and NP type ($\beta = 89.4$, $SE = 54.4$, $z = 1.644$; $p = .1$). However, there are significant interactions of plausibility and syntactic integration score ($\beta = 1316.2$, $SE = 574.9$, $z = 2.289$; $p = .022$), NP type and syntactic integration score ($\beta = 3239.6$, $SE = 1110.8$, $z = 2.916$; $p = .004$), as well as plausibility, NP type, and syntactic integration score ($\beta = 1972.8$, $SE = 703.0$, $z = 2.806$; $p = .005$). In order to establish the nature of the interactions, planned mixed ANOVAs with the factors plausibility and NP type were run for each group. There is a significant effect of plausibility for the low integration group (509 ms for plausible vs. 440 ms for implausible NPs ($\beta = 68.0$, $SE = 31.6$, $z = 2.149$; $p = .035$), yet there is no effect for the mid integration group (444 vs. 418 ms; $z < 1$) or the high integration group (454 vs. 427 ms; $z < 1$). Conversely, for NP type, the low integration group (483 ms for full NPs vs. 454 ms for pronouns; $z < 1$) or the mid integration group (430 vs. 433 ms) do not show effects, whereas the high integration group makes a significant difference according to NP type (463 vs. 392 ms; $\beta = 80.7$, $SE = 33.3$, $z = 2.420$; $p = .017$). This contrast underscores that the low syntactic integration group processes sentences by plausibility, regardless of case marking, whereas the high syntactic integration group parses by case marking, regardless of plausibility differences.

Discussion

Dividing the group according to their syntactic integration scores yield pronouncedly different processing patterns, summarized in Table 10. We discuss the specific processing patterns found for each group in turn.

Low integration group. The reading patterns of the low integration group suggest that this group robustly recruits plausibility information as well as subcategorization information, as shown in the significant effects in the *plausibility* and the *transitivity comparisons.* For the *case comparison*, the significant effect of group in the comparisons with the natives suggests nonnative-like processing of case.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Table 10. *Summary of the findings by group*

		L2 Group ($n = 75$)			
	Natives ($n = 18$) All	All	Low Integration ($n = 25$)	Mid Integration ($n = 25$)	High Integration ($n = 25$)
Garden-path: *control comparison* (8a) & (8b)	Yes	Yes			
Use of subcategorization: *transitivity comparison* (8b) & (8d)	Yes	Yes	Yes	Yes	Yes
Use of morphosyntax: *case comparison* (8a) & (8c)	Yes	No	No	No	Yes
Use of plausibility: *plausibility comparison* (8d) & (8e)	Yes	Yes	Yes	No	Partial

Note: L2, second language.

The low syntactic integrators come back more often to the main clause verb in sentences with nominative-marked pronouns (8c) than in the comma condition in (8a). In addition, the group shows significant differences on the main clause verb in the processing of pronouns depending on whether they could act as plausible objects of the first verb. Even though they are grammatically not accessible as objects by virtue of nominative case marking, the low syntactic integration group distinguishes them according to plausibility. Seen in conjunction with the strong effects in the *plausibility comparison*, these findings suggest that reanalysis in the processing of temporary object–subject ambiguities in the low syntactic integration group is driven by plausibility information but not morphosyntactic information.

Mid integration group. The mid integration group also makes significant contrasts for the *transitivity comparison*. As the low integration group, the group displays similar processing patterns for the *case comparison* with significant differences between sentences in the pronoun condition (8c) and the control condition (8a). These effects indicate that readers initially try to integrate the pronoun as a direct object and need to undo this association. However, this group does not show a *plausibility* effect in that there are no differences between the implausible (8d) and plausible (8e) object conditions on any measures. Note, though, that the mid integration group has the fastest overall reading times (Table 8) of all groups. In their self-paced reading study on similar sentences, Roberts and Felser (2011) also report that faster L1 readers of English do not show a plausibility effects and that faster L2 readers of English showed weaker effects of plausibility.

High integration group. Finally, the high integration group demonstrates nativelike processing patterns in the *transitivity comparison* as well as the *case comparison*. For case, the high integration group shows no differences in the

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

pronoun condition (8c) compared to the comma condition (8a) on the main clause verb or the final region. Further, only NP type and not plausibility affects reading times for pronouns in this group. In consequence, this group appears to have successfully recruited nominative case information on the pronoun for integrating it as the subject of the main clause, such that no reanalysis effort is detectable on the main clause verb region of the sentence. For the *plausibility comparison*, the high integration group also demonstrates differences between implausible and plausible object sentences (8d and 8e) in second-pass measures, although at comparatively weaker levels than natives.

GENERAL DISCUSSION

This study examined whether individual differences modulate the L2 processing of temporary subject–object ambiguities. To this end, the results from an eye tracking study on L2 reading comprehension were correlated with L2 proficiency, working memory capacity, reading speed, automaticity of lexical access, and semantic and syntactic integration ability. The main findings from this study can be summarized as follows:

- For the whole group of 75 L2 readers, reading patterns were affected by subcategorization and plausibility information in measures of second-pass parsing on the main clause verb region. This suggests that the incremental revision of garden paths is affected by subcategorization and plausibility. At the same time, the L2 learners do not appear to integrate case marking incrementally. In this respect, the L2 group taken as a whole differed from the native control group that uses all three types of information, including case marking, for revising garden path sentences.
- In linear mixed-effects regression analysis with individual differences factors, the reading patterns were affected by the automaticity of lexical processing in first-pass measures, the global measure of reading speed for plausibility, and by syntactic integration ability in second-pass measures. Syntactic integration ability significantly correlated with the use of case, subcategorization, and plausibility information in reanalyzing temporary object–subject ambiguities.
- When divided into three groups based on syntactic integration ability, the low integration group showed robust incremental use of plausibility and subcategorization information for recovery from garden paths but no use of morphosyntactic information.
- Like the low integration group, the mid integration group demonstrated use of subcategorization but not morphosyntactic information in on-line sentence revision. Plausibility did not lead to processing differences.
- The high integration group showed a fully nativelike processing pattern in that it recruited morphosyntactic as well as subcategorization information incrementally in first-pass and second-pass measures of processing. The high integration group also showed effects of plausibility, albeit at a lower level than the natives.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

In this section, we discuss these findings against the backdrop of previous research on L2 processing and with respect to its implications for models of L2 processing. The results of the entire group of L2 learners confirm previous findings of studies that looked at the use of particular types of information in isolation. For the whole group, effects of subcategorization (see also Dussias & Scaltz-Cramer, 2008; Frenck-Mestre & Pynte, 1997) and plausibility (see also Roberts & Felser, 2011; Williams, 2006; Williams et al., 2001) influence how fast L2 readers can recover from an initial erroneous parsing commitment. In line with prior work, this study does not find general effects of the incremental integration of case marking in advanced L2 processing (e.g., Hopp, 2006; Jackson & Dussias, 2009). This asymmetry between subcategorization and plausibility information on the one hand and case marking points on the other hand to a greater relative difficulty in accessing morphosyntactic information in L2 processing (e.g., Clahsen & Felser, 2006b; Jiang, 2004; Ullman, 2005).

At the center of the present study is the question whether the use of different types of information in L2 processing is modulated by individual differences. We discuss the influence of these variables in turn.

Although proficiency has been shown to affect the processing of L2 morphosyntax (e.g., Frenck-Mestre, 2002; Hopp, 2006; Rossi et al., 2006; Sagarra & Herschensohn, 2011), it did not exert any measurable influence on the resolution of temporary object–subject ambiguities. Many of the previous studies reporting proficiency effects in L2 morphosyntactic processing show that lower proficient learners do not use morphosyntactic information in processing to the extent that they arrive at an erroneous interpretation of the sentence main clause verb clearly forces undoing the thematic association of the postverbal noun as the object of the first verb, so that all participants reliably converged on the target interpretation of the stimuli. Hence, the reanalysis necessary in this study might not have been challenging enough for learners at high-intermediate to near-native proficiency levels to elicit proficiency-related differences.

Working memory capacity did not have any robust association with any of the reading time effects investigated in this study. The null effect of reading span might be because the sentences in (8) did not pose much of a memory load. Rather, the regions of ambiguity and disambiguation were directly adjacent, which reduces the time competing interpretations of the unfolding sentence would need to be maintained in parallel. Hence, different effects of reading span may be found with sentences that require the storage of constituents in the resolution of ambiguities in long-distance filler-gap dependencies (see Dussias & Pinar, 2010) or nonlocal morphosyntactic agreement relations (e.g., Keating, 2010).

Next, we investigated whether automaticity in basic linguistic processing in word recognition interacts with the use of information in ambiguity resolution. Following the logic in Dekydtspotter et al. (2006) and McDonald (2006, 2010), we assumed that readers whose lexical processing is more effortful run out sooner of resources for integrating grammatical information in the resolution of syntactic ambiguities. Effects of the automaticity of lexical access, as measured in the coefficient of variance, were found in some early measures, which indicate that participants with more automatized lexical access routines show effects of

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

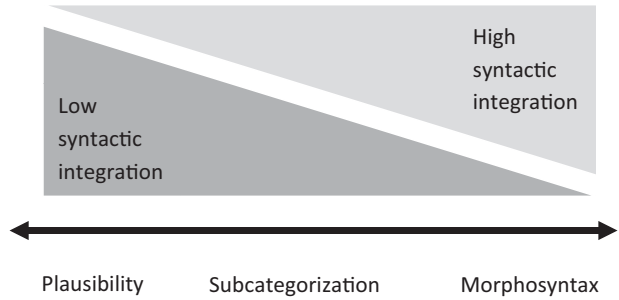


Figure 3. The relation between syntactic integration score and the use of different information types.

transitivity and plausibility earlier on; however, the effects of lexical automaticity did not persist to later reading measures associated with reanalysis and sentence integration.

Overall reading speed turned out to have some marginal associations with the experimental manipulation as regards the use of plausibility information (see also Roberts & Felser, 2011). However, as a global measure of processing efficiency, reading speed was also weakly correlated with almost all other individual differences variables in this study. Hence, it is difficult to pinpoint the nature of its effects and to identify its causes.

One of the underlying causes for individual differences in L2 sentence processing could be the efficiency with which particular types of grammatical information are integrated in context (Kilborn, 1992; Oliver et al., 2012; Sorace, 2005). To establish an independent measure of the integration ability of grammatical information in context, we calculated a score of syntactic integration ability from the normalized speed benefit participants experienced when provided with syntactic information in a word monitoring task, and a score of semantic integration ability when additional semantic information became available. Syntactic and semantic integration scores correlated negatively with each other, indicating that L2 readers exploit different information types. The syntactic integration score also correlated significantly with the main reading effects for the use of morphosyntax, subcategorization, and plausibility information. Subsequent analyses by syntactic integration group showed that all groups make use of subcategorization information. Further, however, they demonstrated distinct processing patterns, with the low integration group capitalizing on plausibility information, yet making no immediate use of case marking information for reanalyzing syntactic ambiguities. In contrast, the high integration group showed robust incremental integration of morphosyntactic information, yet they seemed to be guided less strongly by plausibility information. Putting the respective processing patterns of the three groups in perspective, we can observe a cline in the involvement of different information types according to syntactic integration ability as graphed in Figure 3.

Note that the differences in the information types L2 readers rely on in processing does not seem to incur differences in overall reading speed or comprehension

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

accuracy. As shown in Table 9, the low and the high integration groups were comparably fast in reading the sentences and comparably accurate in comprehending the set of sentences in (8). Instead, individual differences in syntactic integration ability appear to point to two different routes in L2 processing that can be optionally taken, namely, a route relying predominantly on morphosyntactic information and a second route capitalizing on plausibility information.

Such a distinction is reminiscent of how monolingual sentence processing research has characterized native processing in dual-pathway models as the "good enough" model by Ferreira and colleagues (e.g., Ferreira, 2003) and the (late assignment of syntax) sentence processing model by Townsend and Bever (2001). According to these models, sentence processing can proceed along two routes: (a) a shallow analysis that relies on frequent and predictable surface patterns including lexical–semantic and plausibility information; and (b) a deep, syntactic parse that exploits grammatical structure. A host of native processing studies (for review, see Ferreira & Patson, 2007) demonstrates that both routes operate simultaneously in processing in that natives construe a shallow analysis and later revise their heuristic analysis by a deep parse. In the present study, the native controls were found to use lexical–thematic, plausibility, and case information, which suggests that they completed a full grammatical parse.

With respect to L2 processing, several models have adopted a similar track, arguing that sentence processing can proceed along a shallow route, prioritizing surface, lexical–semantic, and pragmatic information, and a deep route that spells out syntactic detail in incremental processing (Clahsen & Felser, 2006b; for related claims, see also Ullman, 2005). L2 learners have been argued to be largely restricted to the shallow route in processing the L2, with the deep route being limited, if available at all, to the most highly proficient L2 speakers (Clahsen & Felser, 2006a). The present study is broadly compatible with the basic assumptions of these L2 processing models; however, it shows, first, that shallow processing does not characterize L2 processing in general and, second, that proficiency may not be the crucial factor in determining the involvement of morphosyntax in L2 processing (for similar results from event-related potential studies, see also Tanner, 2011; Tanner, McLaughlin, Herschensohn, & Osterhout, in press). Rather, the multifactorial comparisons based on a large group of participants in this study identify individual differences in syntactic integration ability as arbitrating the use of morphosyntactic information versus plausibility information.[5]

These findings suggest that we need to reassess the interpretation of the data from previous studies. Contrary to their suggestions, shallow processing may well turn out not to be a characteristic of (advanced) L2 processing in general; rather, it may characterize a particular subset of L2 participants who adopt the "shallow" route. Nevertheless, these learners may be the predominant type of L2 learner among the general set of advanced L2 speakers. In the analysis of the entire group of 75 participants, plausibility and subcategorization information were found to guide recovery from garden paths, whereas case information on the pronoun was not. At the group level of a larger set of advanced L2 speakers who are not tested for individual differences, then, shallow processing may seem characteristic of nonnative language processing in general. At the same time, group results mask significant and systematic individual differences.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Once the effects of individual differences in syntactic integration ability are taken into account, we can identify a group of adult L2 learners that shows nativelike processing of morphosyntax. For the high syntactic integration group, the mean syntactic integration score was statistically indistinguishable from the score of the native group, F (1, 41) = 3.516, p = .068, with the L2 group having a numerically higher score (M = 0.112) than the natives (M = 0.086). In contrast, the low integration group had a significantly lower syntactic integration score (M = 0.011) than the natives, F (1, 41) = 28.970, p < .001. These differences translate into different reading patterns of ambiguous sentences in that the high integration group demonstrates nativelike processing of morphosyntax, whereas the low integration group is nonnative-like with respect to morphosyntax, yet it demonstrates strong effects of plausibility. These systematic relations between the independent measure of integration ability and on-line sentence processing indicate that L2 processing can occur along two different routes.

These findings immediately raise questions about the nature of these two routes. On the basis of the present findings, the choice of a particular route does not seem to derive from proficiency, constraints in working memory resources or problems in lower level linguistic processing. It would be desirable to explore whether integration ability is related to other measures of individual differences, such as language aptitude (e.g., Robinson, 2005), cognitive control (e.g., Hachmann, Konieczny, & Müller, 2009; Novick, Trueswell, & Thompson-Schill, 2005), or other measures of linguistic processing, for example, phonological decoding ability (McDonald, 2006). However, since aptitude typically correlates very robustly with proficiency scores (e.g., Dörnyei & Skehan, 2003), and phonological decoding ability directly feeds into lexical recognition, we might have at least observed some indirect effects of those factors in the present study.

Hence, differences in integration ability may index the general availability of two optional routes in L2 processing much the same way these two routes are available in parallel in monolingual processing. For native parsing, Ferreira and Patson (2007) argue that the coexistence of two routes is characteristic of a system that trades off a quick heuristics for understanding sentences and a thorough analysis of sentences. Ferreira, Engelhardt, and Jones (2009) conceptualize good-enough processing as a consequence of readers adopting a so-called satisficing approach to the task at hand, that is, they optimize processing with a view to making a task-relevant decision. In a visual world eye tracking experiment, they found that listeners sacrificed accuracy in computing garden path sentences for accuracy in carrying out subsequent commands. In addition, Ferreira and colleagues found that increasing task demands decreases attention to linguistic detail. In other studies, varying the format of the task affects the use of plausibility information in processing reduced relative clause ambiguities (Long & Prat, 2008). Swets, Desmet, Clifton, and Ferreira (2008) report compatible findings from self-paced reading of ambiguous sentences, where the type of comprehension question affected the degree to which grammatical and extragrammatical information was accessed in resolving the ambiguity in processing.

In the present study, participants faced the task of answering the comprehension questions following each experimental sentence. Given the high comprehension

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

accuracy across the board, participants seem to have focused on completing this aspect of the overall task. In order to achieve high comprehension accuracy, the L2 learners likely relied on integrating the information types they were most adept at utilizing. Therefore, the individual differences found in this study may be interpreted as illustrating how L2 readers allocate their resources to the best of their capacities for completing the task at hand. Rather than being due to lack of grammatical knowledge or fixed capacity limitations, individual differences in the use of morphosyntax thus may reflect the optimal mapping of different types of information to a sentence interpretation under the specific experimental conditions.

In similar terms, Wilson, Sorace, and Keller (2009) describe patterns in the L2 acquisition of discourse-related phenomena in terms of a resource limitation and resource allocation approach. It is argued that L2 learners need to allocate their resources in L2 processing efficiently, because they are, overall, more limited in memory and processing resources in the L2 than in the L1. As a consequence, L2 learners may prioritize one kind of information at the expense of others. The present results are compatible with this approach and suggest that L2 processing involves a complex interaction of learner characteristics, sentence properties, and task effects.

Although this study is but a starting point in the systematic investigation of individual differences in L2 processing, it highlights the need for more research examining whether individual differences in syntactic integration ability affect the L2 processing of other phenomena, in particular long-distance dependencies for which it has been claimed that L2 speakers do not recruit abstract syntactic information (Clahsen & Felser, 2006b; Marinis et al., 2005; but see Pliatsikas & Marinis, 2013). In addition, it may be fruitful to investigate whether the different processing strategies can be affected by task demands, changes in ancillary tasks, or differences in L2 experience. For instance, repeated exposure to or training on particular structures may increase use of morphosyntactic information (e.g., Long & Prat, 2008). Ultimately, it will be essential to understand the scope of individual differences in L2 sentence comprehension before we can make generalizations about the use of respective information types or processing routes in L2 processing.

APPENDIX A

This appendix lists all experimental items of the reading experiment. The first verb of the adjunct clause is used in the control condition and the intransitive condition, the second in the optionally transitive conditions. The first postverbal NP is used in the control, the intransitive, and the implausible conditions. The pronoun is used in the pronoun condition. The letter in parentheses marks whether the pronoun was judged as a plausible (P) or implausible (I) object of the preceding optionally transitive verb. The third NP was used in the plausible condition.

1. When the girl was praying/playing the boy/he (I)/the piano made some funny noises.
2. When the girl screamed/drove the mouse/he (P)/the car moved away from the cat.
3. As the audience was clapping/applauding the theatre/they (P)/the actor shook and trembled.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

4. When the boy appeared/drank the girl/she (I)/the milk got angry/cold in the kitchen.
5. While the students listened/sang the teacher/he (I)/the tune forced them to pay attention.
6. When the girl sneezed/called the boy/he (P)/the phone fell onto the floor.
7. As the man talked/asked the truck/she (P)/the woman made some phone calls/weird noise.
8. When the baby smiled/ate the nurse/she (I)/the pie started getting happy/cold.
9. When the girl danced/walked the mother/she (I)/the dog found some coins.
10. When Sue fell/dressed the music/she (P)/the girl stopped and helped her/continued later.
11. When the doctor laughed/read the woman/she (I)/the book lay in his lap.
12. When the cat was sleeping/climbing the bird/they (P)/the tree lost some leaves/feathers.
13. Because the man was snoring/hiding the customers/they (P)/the jewels moved/disappeared to the shop next door.
14. Before the pilot coughed/landed the woman/she (I)/the plane read some magazines/lost some speed.
15. As the dog was sleeping/fighting the carrot/she (P)/the parrot moved around on the couch/counter.
16. When the customer complained/parked the salesman/they (P)/the limousine left for the boss/hotel.
17. When the plane was rising/leaving the child/she (I)/the airport closed the eyes/gate.
18. As the boy was hopping/racing the girl/she (I)/the bike smashed into a car.
19. While the children were giggling/humming the adults/they (I)/the song went/broke off at some point.
20. Because the woman was sobbing/baking the man/he (I)/the cake needed some ice cream/icing.
21. After the girl came/cleaned the woman/she (P)/the kitchen looked more beautiful.
22. When the captain died/sailed the crew/they (I)/the boat made some turns to the coast.
23. While the aunt lived/entered the nephew/he (I)/the building looked after her/collapsed around her.
24. When the horse was leaping/running the cow/she (I)/the race became more agitated.
25. While the woman was bowing/spinning the crowd/they (I)/the yarn got excited/twisted by the applause/into a knot.
26. As the teacher disappeared/advised the lesson/they (P)/the students became more interesting/interested.
27. When the dog was barking/chasing the television/they (P)/the children started smoking at the top/in the street.
28. As the lion was growling/hunting the rain/they (P)/the zebra began running away/pouring down.
29. When the woman arrived/phoned the train/he (P)/the man broke down with a heart/engine failure.
30. Because the actor lied/phoned the curtain/they (P)/the theatre remained closed for the evening.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Table A.1. *Mean reading times (ms) and number of regressions for the low syntactic integration group (n = 25)*

		Adjunct Clause Verb	Noun Phrase	Main Clause Verb	Final Region
(8a) Comma	First pass	331 (18)	325 (14)	290 (14)	688 (31)
	Second pass	271 (24)	309 (24)	291 (37)	507 (46)
	Total	424 (23)	443 (21)	377 (23)	827 (33)
	Regressions	29	18	19	104
(8b) Intransitive	First pass	280 (13)	400 (17)	307 (15)	722 (34)
	Second pass	336 (33)	357 (49)	304 (35)	468 (59)
	Total	404 (24)	504 (25)	404 (25)	852 (40)
	Regressions	29	25	17	84
(8c) Pronoun	First pass	268 (11)	255 (11)	289 (15)	670 (34)
	Second pass	386 (40)	347 (50)	364 (32)	526 (47)
	Total	455 (30)	398 (29)	459 (27)	908 (42)
	Regressions	36	15	39	114
(8d) Implausible	First pass	266 (14)	404 (18)	287 (12)	684 (31)
	Second pass	342 (25)	437 (35)	342 (31)	479 (58)
	Total	429 (31)	587 (29)	426 (22)	841 (35)
	Regressions	30	26	28	102
(8e) Plausible	First pass	281 (13)	393 (20)	327 (16)	650 (33)
	Second pass	389 (30)	472 (42)	415 (30)	586 (47)
	Total	476 (26)	674 (38)	541 (28)	898 (39)
	Regressions	32	44	46	128

Note: The regions of interest are delimited by brackets: When the girl [was playing] [the boy] [made] [some funny noises]. Standard errors are given in parentheses.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Table A.2. *Mean reading times (ms) and number of regressions for the mid syntactic integration group (n = 25)*

		Adjunct Clause Verb	Noun Phrase	Main Clause Verb	Final Region
(8a) Comma	First pass	271 (14)	314 (13)	270 (11)	658 (36)
	Second pass	285 (28)	341 (44)	303 (33)	486 (83)
	Total	371 (20)	401 (22)	360 (20)	748 (43)
	Regressions	38	13	31	84
(8b) Intransitive	First pass	265 (12)	375 (17)	273 (12)	643 (36)
	Second pass	390 (47)	359 (36)	309 (30)	542 (66)
	Total	405 (32)	510 (26)	366 (19)	756 (43)
	Regressions	24	30	24	75
(8c) Pronoun	First pass	253 (10)	256 (14)	311 (16)	653 (39)
	Second pass	331 (37)	266 (28)	358 (44)	476 (52)
	Total	390 (23)	345 (23)	433 (26)	817 (51)
	Regressions	23	10	46	89
(8d) Implausible	First pass	264 (12)	370 (15)	296 (15)	667 (40)
	Second pass	372 (37)	465 (51)	329 (30)	474 (69)
	Total	413 (25)	570 (36)	428 (24)	806 (45)
	Regressions	23	30	35	91
(8e) Plausible	First pass	254 (12)	362 (14)	300 (13)	626 (34)
	Second pass	299 (26)	416 (43)	333 (42)	618 (72)
	Total	370 (19)	536 (29)	431 (27)	834 (48)
	Regressions	33	25	30	109

Note: The regions of interest are delimited by brackets: When the girl [was playing] [the boy] [made] [some funny noises]. Standard errors are given in parentheses.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Table A.3. *Mean reading times (ms) and number of regressions for the high syntactic integration group (n = 25)*

		Adjunct Clause Verb	Noun Phrase	Main Clause Verb	Final Region
(8a) Comma	First pass	267 (14)	287 (13)	269 (13)	642 (34)
	Second pass	266 (27)	328 (30)	322 (46)	384 (49)
	Total	376 (19)	404 (22)	381 (24)	773 (39)
	Regressions	35	27	26	93
(8b) Intransitive	First pass	256 (11)	362 (18)	263 (13)	591 (29)
	Second pass	370 (36)	376 (29)	280 (21)	556 (54)
	Total	392 (24)	523 (26)	347 (17)	741 (35)
	Regressions	22	48	21	97
(8c) Pronoun	First pass	239 (12)	258 (13)	290 (16)	684 (39)
	Second pass	355 (38)	269 (23)	295 (26)	432 (62)
	Total	408 (25)	344 (19)	392 (20)	865 (54)
	Regressions	29	29	31	102
(8d) Implausible	First pass	242 (11)	337 (16)	311 (14)	655 (34)
	Second pass	310 (23)	398 (30)	296 (26)	449 (39)
	Total	404 (21)	550 (30)	429 (21)	803 (39)
	Regressions	35	45	37	99
(8e) Plausible	First pass	240 (11)	353 (17)	301 (15)	586 (35)
	Second pass	362 (31)	439 (39)	400 (31)	615 (57)
	Total	423 (25)	582 (35)	499 (27)	871 (48)
	Regressions	40	47	43	123

Note: The regions of interest are delimited by brackets: When the girl [was playing] [the boy] [made] [some funny noises]. Standard errors are given in parentheses.

ACKNOWLEDGMENTS

Parts of this paper were presented at the Workshop on Bilingualism: Neurolinguistic and Psycholinguistic Perspectives in Aix-en-Provence, EUROSLA 22 in Poznan, and GALANA 5 in Lawrence, Kansas. I thank the audiences, Bonnie D. Schwartz, Giuli Dussias, Carrie N. Jackson, as well as three anonymous reviewers for helpful comments on earlier versions of the manuscript. Thanks also go to Lena Holderer for assisting with data collection. All remaining shortcomings are my responsibility.

NOTES

1. However, L2 learners are found to be more sensitive to inflection if they perform ancillary judgment or comprehension tasks that relate to the experimental manipulation of morphosyntax (Havik et al., 2009; Jackson & Dussias, 2009).
2. An anonymous reviewer wonders whether it would have been more appropriate to test working memory in the L1 rather than in the L2. This study follows all previous experiments on working memory effects on L2 processing in testing reading span in the L2. Studies comparing L1 and L2 reading spans report moderate to high positive correlations (e.g., Alptekin & Erçetin, 2010; Miyake & Friedman, 1998; Service, Simola, Metsänheimoi, & Maury, 2002), so that assessing working memory exclusively in the L2 seems valid. Another concern is that the reading span task may have been too demanding because it included a judgment as well as a recall task. Mean judgment accuracy was 84% in the natives ($M = 59/70$, $SD = 5.9$) and 67% in the L2 group ($M = 47$, $SD = 7.7$). For the reading span score used in this study, we calculated both a composite score of recall and judgment accuracy (as in Ariji et al., 2003) and a score based only on recall accuracy. For the L2 learners, these scores are highly correlated ($r = .771$, $p < .001$), as they are for the natives ($r = .920$, $p < .001$). Similarly, the composite reading span score and judgment accuracy are highly correlated (L2 group: $r = .733$; L1 group: $r = .778$).
3. It was decided to compare the pronoun condition with the full NP conditions in the control condition (8a), rather than, for example, a condition with postverbal pronouns that are ambiguous in case marking (*it*). Given the difference in animacy between *s/he* and *it*, the change of pronoun would entail a difference in plausibility for most sentences (e.g., *When the boy drank he/it*). The difference in case marking would hence be confounded with a difference in plausibility.
4. This study only explores individual differences in L2 processing. Although data on individual differences were also collected from the L1 participants, we do not present them because, first, the group size is too small for a meaningful exploration of individual differences and, second, the L1 group differs significantly on all means in reading span, lexical decision times, and integration scores from the L2 group. In all respects, natives are faster and they have higher reading span scores than L2 speakers. Hence, natives and nonnatives are not directly comparable in how individual differences may affect sentence processing.
5. However, all participants in this study came from a fairly homogeneous student population, and the proficiency scores ranged from high intermediate to near native, with almost 90% of all participants scoring in the advanced and higher range as per the Oxford Placement Test (Allen, 1992). It is hence conceivable that larger differences in proficiency may be associated with the different routes observed in L2 processing,

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

and future research should investigate this matter. Nevertheless, almost all previous studies on L2 sentence processing recruited participants from comparable populations assessed by the same proficiency test and report that these participants as a group do not project morphosyntactic detail in on-line comprehension of the L2 (e.g., Felser & Roberts, 2007; Marinis et al., 2005).

REFERENCES

Adams, B. C., Clifton, C., Jr., & Mitchell, D. C. (1998). Lexical guidance in sentence processing? *Psychonomic Bulletin & Review, 5*, 265–270.

Allan, D. (1992). *The Oxford Placement Test*. Oxford: Oxford University Press.

Alptekin, C., & Erçetin, G. (2010). The role of L1 and L2 working memory in literal and inferential comprehension in L2 reading. *Journal of Research in Reading, 33*, 206–219.

Ariji, K., Omaki, A., & Tatsuta, N. (2003). Working memory restricts the use of semantic information in ambiguity resolution. In P. Slezak (Ed.), *Proceedings of the 4th International Conference on Cognitive Science* (pp. 19–25). Sydney: University of New South Wales.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412.

Bowden, H. W., Sanz, C., & Stafford, C. A. (2005). Individual differences: Age, sex, working memory, and prior knowledge. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp. 105–140). Washington, DC: Georgetown University Press.

Clahsen, H., & Felser, C. (2006a). Continuity and shallow structures in language processing. *Applied Psycholinguistics, 27*, 107–126.

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

Clifton, C., Jr. (1993). Thematic roles in sentence parsing. *Canadian Journal of Experimental Psychology, 47*, 222–246.

Coderre, E., van Heuven, W. J. B., & Conklin, K. (2011.) *Lexical processing is delayed by 100ms in a second language.* Paper presented at the Workshop on Bilingualism: Neurolinguistic and Psycholinguistic Perspectives, Aix en Provence.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Jambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and a user's guide. *Psychonomic Bulletin and Review, 12*, 769–786.

DeDe, G. (2010). Utilization of prosodic information in syntactic ambiguity resolution. *Journal of Psycholinguistic Research, 39*, 345–374.

Dekydtspotter, L., Schwartz, B. D., & Sprouse, R. A. (2006). The comparative fallacy in L2 processing research. In M. Grantham O'Brien, C. Shea, & J. Archibald (Eds.), *Proceedings of the 8th Generative Approaches to Second Language Acquisition Conference (GASLA 2006): The Banff Conference* (pp. 33–40). Somerville, MA: Cascadilla Press.

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

Dörnyei, Z., and Skehan, P. (2003). Individual differences in second language learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 589–630). Oxford: Blackwell.

Dussias, P. E., & Cramer Scaltz, T. R. (2008). Spanish–English L2 speakers' use of subcategorization bias information in the resolution of temporary ambiguity during second language reading. *Acta Psychologica, 128*, 501–513.

Dussias, P. E., & Pinar, P. (2010). Effects of reading span and plausibility in the reanalysis of wh-gaps by Chinese–English second language speakers. *Second Language Research, 26*, 443–472.

Farmer, T. A., Misyak, J. B., & Christiansen, M. H. (in press). Individual differences in sentence processing. In M. J. Spivey, K. McRae, & M. Joanisse (Eds.), *The Cambridge handbook of psycholinguistics*. New York: Cambridge University Press.

Felser, C., & Roberts, L. (2007). Processing wh-dependencies in a second language: A cross-modal priming study. *Second Language Research, 23*, 9–36.

Felser, C., Roberts, L., Marinis, T., & Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics, 24*, 453–489.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*, 164–203.
Ferreira, F., Engelhardt, P. E., & Jones, M. W. (2009). Good enough language processing: A satisficing approach. In N. Taatgen, H. Rijn, J. Nerbonne, & L. Schomaker (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 413–418). Austin, TX: Cognitive Science Society.
Ferreira, F., & Patson, N. D. (2007). The "good enough" approach to language comprehension. *Language and Linguistics Compass, 1*, 71–83.
Frazier, L., & Clifton, C. (1998). Sentence reanalysis and visibility. In J. D. Fodor & F. Ferreira (Eds.), *Reanalysis in sentence processing*. Dordrecht: Kluwer.
Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In R. R. Heredia & J. Altarriba (Eds.), *Bilingual sentence processing* (pp. 217–236). Amsterdam: Elsevier.
Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology, 50A*, 119–148.
Goethe Institut. (2010). *German Placement Test.* Retrieved October 2012 from http://www.goethe.de/cgi-bin/einstufungstest/einstufungstest.pl
Hachmann, W., Konieczny, L., & Müller, D. (2009). Individual differences in the processing of complex sentences. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 309–314). Austin, TX: Cognitive Science Society.
Harrington, M. (2007). The coefficient of variance as an index of L2 lexical processing skill. *University of Queensland Working Papers in Linguistics, 1*, 2–43.
Havik, E., Roberts, L., van Hout, R., Schreuder, R., & Haverkort, M. (2009). Processing subject–object ambiguitities in the L2: A self-paced reading study with German L2 learners of Dutch. *Language Learning, 59*, 73–112.
Holmes, V., Stowe, L., & Cupples, L. (1989). Lexical expectations in parsing complement–verb sentences. *Journal of Memory and Language, 28*, 668–689.
Hopp, H. (2006). Syntactic features and reanalysis in near-native processing. *Second Language Research, 22*, 369–397.
Hopp, H. (2007). *Ultimate attainment at the interfaces in second language acquisition: Grammar and processing*. Groningen: Grodil Press.
Hopp, H. (2009). The syntax–discourse interface in near-native L2 acquisition: Off-line and on-line performance. *Bilingualism: Language and Cognition, 12*, 463–483.
Hopp, H. (2010). Ultimate attainment in L2 inflectional morphology: Performance similarities between non-native and native speakers. *Lingua, 120*, 901–931.
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.
Jackson, C. N., & Bobb, S. C. (2009). The processing and comprehension of wh-questions among second language speakers of German. *Applied Psycholinguistics, 30*, 603–636.
Jackson, C. N., & Dussias, P. E. (2009). Cross-linguistic differences and their impact on L2 sentence processing. *Bilingualism: Language and Cognition, 12*, 65–82.
Jegerski, J. (2012). The processing of temporary subject–object ambiguities in native and near-native Mexican Spanish. *Bilingualism: Language and Cognition, 15*, 721–735.
Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25*, 603–634.
Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning, 57*, 1–33.
Juffs, A. (2004). Representation, processing, and working memory in a second language. *Transactions of the Philological Society, 102*, 199–225.
Juffs, A., & Harrington, M. (1996). Garden path sentences and error data in second language processing research. *Language Learning, 46*, 286–324.
Keating, G. D. (2010). The effects of linear distance and working memory on the processing of gender agreement in Spanish. In B. Van Patten and J. Jegerski (Eds.), *Research in second language processing and parsing* (pp. 113–134). Amsterdam: John Benjamins.
Kilborn, K. (1992). On-line integration of grammatical information in a second language. In R. Harris (Ed.), *Cognitive processing in bilinguals* (pp. 337–350). Amsterdam: Elsevier.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Liu, Y., & Perfetti, C. A. (2003). The time course of brain activity in reading English and Chinese: An ERP study of Chinese bilinguals. *Journal of Human Brain Mapping, 18*, 167–175.
Long, D. L., & Prat, C. S. (2008). Individual differences in syntactic ambiguity resolution: Readers vary in their use of plausibility information. *Memory and Cognition, 36*, 375–391.
Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27*, 53–78.
McDonald, J. L. (2006). Beyond the critical period: Processing-based explanations for poor grammaticality judgment performance by late second language learners. *Journal of Memory and Language, 55*, 381–401.
McDonald, J. L., & Roussel, C. C. (2010). Past tense grammaticality judgment and production in non-native and stressed native English speakers. *Bilingualism: Language and Cognition, 13*, 429–448.
McLaughlin, J., Tanner, D., Pitkänen, I, Frenck-Mestre, C., Inoue, K., Valentine, G., et al. (2010). Brain potentials reveal discrete stages of L2 grammatical learning. *Language Learning, 60*(Suppl. 2), 123–150.
Michael, E., & Gollan, T. (2005). Being and becoming bilingual: Individual differences and consequences for language production. In J. Kroll & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 389–410). Oxford: Oxford University Press.
Mitchell, D. C. (1994). Sentence parsing. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 375–409). San Diego, CA: Academic Press.
Miyake, A., & Friedman, N. (1998). Individual differences in second language proficiency: Working memory as language aptitude. In A. F. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 339–364). Mahwah, NJ: Erlbaum.
Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience, 5*, 263–281.
Oliver, G., Gullberg, M., Hellwig, F., Mitterer, H., & Indefrey, P. (2012). Acquiring L2 sentence comprehension: A longitudinal study of word monitoring in noise. *Bilingualism: Language and Cognition, 15*, 841–857.
Omaki, A., & Schulz, B. (2011). Filler-gap dependencies and island constraints in second language sentence processing. *Studies in Second Language Acquisition, 33*, 563–588.
Pickering, M. J., & Traxler, M. J. (1998). Plausibility and recovery from garden paths: An eye-tracking study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*, 940–961.
Pickering, M. J., Traxler, M. J., & Crocker, M. W. (2000). Ambiguity resolution in sentence processing: Evidence against frequency-based accounts. *Journal of Memory and Language, 43*, 447–475.
Pliatsikas, C., & Marinis, T. (2013). Processing empty categories in a second language: When naturalistic exposure fills the (intermediate) gap. *Bilingualism: Language and Cognition, 16*, 167–182.
Pritchett, B. L. (1992). *Grammatical competence and parsing performance*. Chicago: University of Chicago Press.
Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*, 372–422.
Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences, 26*, 445–526.
Roberts, L., & Felser, C. (2011). Plausibility and recovery from garden paths in second language sentence processing. *Applied Psycholinguistics, 32*, 299–331.
Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics, 25*, 46–73.
Rossi, S., Gugler, M. F., Friederici, A. D., & Hahne, A. (2006). The impact of proficiency on syntactic second-language processing of German and Italian: Evidence from event-related potentials. *Journal of Cognitive Neuroscience, 18*, 2030–2048.
Sagarra, N., & Herschensohn, J. (2011) Proficiency and animacy effects on L2 gender agreement processes during comprehension. *Language Learning, 61*, 80–116.
Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh, PA: Psychology Software Tools.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Segalowitz, N., & Hulstijn, J. (2005). Automaticity in bilingualism and second language learning. In J. Kroll & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371–388). Oxford: Oxford University Press.

Segalowitz, N., & Segalowitz, S. J. (1993). Skilled performance practice and differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics, 13*, 369–385.

Segalowitz, N., Segalowitz, S. J., & Wood, A. G. (1998). Assessing the development of automaticity in second language word recognition. *Applied Psycholinguistics, 19*, 53–67.

Service, E., Simola, M., Metsänheimoi, O., & Maury, S. (2002). Bilingual working memory is affected by language skill. *European Journal of Cognitive Psychology, 14*, 383–408.

Sorace, A. (2005). Syntactic optionality at interfaces. In L. Cornips & K. Corrigan (Eds.), *Syntax and variation: Reconciling the biological and the social* (pp. 46–111). Amsterdam: John Benjamins.

Sorace, A. (2011). Pinning down the concept of "interface" in bilingualism. *Linguistic Approaches to Bilingualism, 1*, 1–33.

Spivey, M. J., & Tanenhaus, M. K. (1998). Syntactic ambiguity resolution in discourse: Modeling the effects of referential context and lexical frequency. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 24*, 1521–1543.

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., White, E., & Drury, J. E. (2009). Temporal dynamics of late second language acquisition: Evidence from event-related potentials. *Second Language Research, 25*, 13–41.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory and Cognition, 36*, 201–216.

Swets, B., Desmet, T., Hambrick, D. Z., & Ferreira, F. (2007). The role of working memory in syntactic ambiguity resolution: A psychometric approach. *Journal of Experimental Psychology: General, 136*, 64–81.

Tanner, D. (2011). *Agreement mechanisms in native and nonnative language processing: Electrophysiological correlates of complexity and interference*. Unpublished doctoral dissertation, University of Washington.

Tanner, D., McLaughlin, J., Herschensohn, J., & Osterhout, L. (in press). Individual differences reveal stages of L2 acquisition: ERP evidence. *Bilingualism: Language and Cognition*.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

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

Trueswell, J., Tanenhaus, M. K., & Kello, C. (1993). Verb specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 528–533.

Ullman, M. T. (2005). A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In C. Sanz (Ed.), *Mind and context in adult second language acquisition: Methods, theory, and practice* (pp. 141–178). Washington, DC: Georgetown University Press.

van Gompel, R. P. G., & Pickering, M. J. (2001). Lexical guidance in sentence processing: A note on Adams, Clifton, and Mitchell (1998). *Psychonomic Bulletin & Review, 8*, 851–857.

van Gompel, R. P. G., Pickering, M. J., Pearson, J., & Liversedge, S. P. (2005). Evidence against competition during syntactic ambiguity resolution. *Journal of Memory and Language, 52*, 284–307.

Van Overschelde, J. P., Rawson, K. A., & Dunlosky, J. (2004). Category norms: An update and expanded version of the Battig and Montague (1969) norms. *Journal of Memory and Language, 50*, 289–335.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press

Vasishth, S., & Drenhaus, H. (2011). Locality in German. *Dialogue and Discourse, 1*, 59–82.
Williams, J. (2006). Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition, 9*, 71–88.
Williams, J., Möbius, P., & Kim, C. (2001). Native and non-native processing of English wh-questions: Parsing strategies and plausibility constraints. *Applied Psycholinguistics, 22*, 509–540.
Wilson, M. P., & Garnsey, S. M. (2009). Making simple sentences hard: Verb bias effects in simple direct object sentences. *Journal of Memory and Language, 60*, 368–392.
Wilson, F., Sorace, A., & Keller, F. (2009) *Simulating L2 learners' deficits at the syntax–discourse interface in native speakers*. Talk presented at the International Symposium on Bilingualism 7, Utrecht.

https://doi.org/10.1017/S0142716413000180 Published online by Cambridge University Press