Applied Psycholinguistics **27** (2006), 107–126
Printed in the United States of America
DOI: 10.1017.S0142716406060206

AUTHORS' RESPONSE

Continuity and shallow structures in language processing

HARALD CLAHSEN and CLAUDIA FELSER
University of Essex

ADDRESS FOR CORRESPONDENCE
Harald Clahsen, Department of Linguistics, University of Essex, Colchester CO4 3SQ, UK.
E-mail: harald@essex.ac.uk.

The core idea that we argued for in the target article was that grammatical processing in a second language (L2) is fundamentally different from grammatical processing in one's native (first) language (L1). Our major source of evidence for this claim comes from experimental psycholinguistic studies investigating morphological and syntactic processing in child and adult native speakers, and nonnative speakers who acquired their L2 after childhood and for whom their L1 is the dominant language. With respect to child L1 processing, we argued for a *continuity of parsing hypothesis* claiming that the child's structural parser is basically the same as that of mature speakers and does not change over time. Adult L2 learners, in contrast, were seen to underuse syntactic information during sentence processing and to rely more on lexical–semantic cues to interpretation. To account for the observed L1/L2 differences in processing, we proposed the shallow structure hypothesis (SSH) according to which the representations adult L2 learners compute during processing contain less syntactic detail than those of child and adult native speakers.

Scholars representing different fields of research commented on our target article, and we thank all of them for their stimulating ideas and detailed criticism. We take the responses we received as a clear sign that the study of grammatical processing in child and adult learners has entered a new and exciting phase. Until fairly recently, grammatical processing in language learners had been subject to much speculation but little empirical investigation of the real-time processes involved in production and comprehension. This has particularly been the case for L2 research, prompting Juffs (2001, p. 207ff) to remark that it is "embarrassing" for the L2 acquisition community that reaction time measures have hardly been used in mainstream L2 acquisition research even though such measures and experimental designs have been available from psychometric experiments with native speakers for nearly a century. We fully agree with Juffs on this point, and also think that the view held by some applied linguists that L1 processing research has little if anything to contribute to the study of L2 (see, e.g., VanPatten, 2004) is misguided. The traditional research gap between experimental psycholinguistics

© 2006 Cambridge University Press 0142-7164/06 $12.00

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

and L2 research has begun to narrow in recent years, however, as witnessed by the research represented in this special issue. The study of language processing in language learners has become a truly interdisciplinary enterprise with a strong cross-linguistic focus, with researchers from linguistics, psychology, and cognitive neuroscience examining different groups of language learners using current behavioral and brain-related psycholinguistic techniques. Although many fundamental questions are still unresolved and further experimental study is clearly required, the picture that has emerged thus far suggests that there are differences between L1 and L2 processing. The question, however, of what the sources of these differences might be remains controversial.

Against this background, the purpose of the target article was to present an overview of empirical findings on grammatical processing in two groups of language learners (child L1 and adult L2), and an attempt at understanding the nature of the observed differences. As will become clear from the discussion below, some commentaries led us to rethink and modify some of the claims we made in our target article, whereas other commentators misinterpreted some of the findings, or indicated a need for clarification. However, we think that the major claim of our target article, that the differences between native and nonnative grammatical processing are real and fundamental, can nevertheless be maintained. The subsequent discussion will be structured as follows. We will first discuss the comments we received on our studies of child L1 processing. We will then elaborate the SSH in response to questions and criticisms raised by several commentators, before addressing some more general issues. Finally, we will make some remarks on potential implications of L2 processing research for foreign language teaching.

CONTINUITY IN CHILDREN'S GRAMMATICAL PROCESSING

Our results on child L1 processing, we argued, are consistent with the continuity hypothesis according to which the child's processing mechanisms (at least in the age range tested) are the same as in mature adults and do not undergo any developmental changes. The dual-mechanism system for processing morphologically complex words (comprising both lexical storage and morphological decomposition) also appears to be available to children, and the observed child/adult differences were argued to result from children's smaller lexicon and slower lexical retrieval. For sentence processing, we found that children apply the same kind of phrase structure based parsing mechanisms as adults during ambiguity resolution, and that children do not differ from adults in their ability to establish syntactic dependencies during on-line sentence comprehension. Differences between children and mature L1 speakers, such as the finding that children rely less on lexical–semantic cues for ambiguity resolution than adults, were argued to result from children's relatively limited short-term memory capacity.

These claims were much less controversial among our commentators than were our claims about adult L2 processing. Avrutin draws attention to off-line interpretation tasks in which children were found to perform differently from adults. Indefrey criticizes our dual-mechanism account of the participle production data (Clahsen, Hadler, & Weyerts, 2004) for not being explicit enough in terms of the *timing* of the different processing mechanisms, and McKee, Rispoli, McDaniel, and Garrett offer a developmental interpretation of these data. Steinhauer raises

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

concerns against how we interpreted the results of the child event-related potential (ERP) study (Lück, Hahne, Friederici, & Clahsen, 2001), and Traxler, although broadly agreeing with our views on child processing, asks for further justification for attributing child/adult processing differences to limitations in lexical knowledge and cognitive resources such as working memory (WM) capacity.

The interpretation data Avrutin mentions are interesting, and children's relatively low response accuracy scores in these off-line experiments indicate a level of confusion that is not seen in adults. Avrutin thinks that this is due to children relying less on syntactic cues for interpretation than adults. But there may be other reasons. Note, for example, that correct picture pointing for so-called D-linked *wh*- questions (*Which lion did the tiger chase*?) requires maintaining two referents in memory (lion, tiger) and selecting among two identical referents (two lions), whereas for non-D-linked *wh*- questions (*Who did the tiger chase*?), only one referent needs to be kept in memory (tiger), which provides a unique reference point for selecting the correct response. Thus, the children's relatively low performance on *which* questions in this task could be a result of cognitive overload (having to deal with too many referents at the same time) rather than an indication of weaker syntactic mechanisms in children than in adults. A similar reasoning may also apply to the exceptional case marking (ECM) data that Avrutin mentions, where an ECM construction introduces extra referents that are not present in the simple transitive control sentences. It should also be noted that as the children in the studies Avrutin refers to were quite young, it is conceivable that they had not yet fully acquired the syntactic knowledge necessary for performing well in their experiments. Where this is the case, children may rely on other sources of information for processing and interpretation. This would not be indicative of a different parser in children, though, but rather a consequence of lack of linguistic knowledge. Hence, the evidence Avrutin reports does not make a case against the continuity hypothesis for children's sentence processing.

Two commentaries (Indefrey and McKee, Rispoli, McDaniel, & Garrett) dealt with the speeded production task we used to investigate processes involved in children's on-line production of inflected words. Indefrey argues that the dual-mechanism account of the reversed frequency effect we obtained entails that blocking (as required for the correct production of irregulars) is made impossible, because the production of unblocked regulars is said to be faster than the retrieval of stored forms. He also notes that this account only applies to specific stimulus sets that contain large numbers of irregulars, such as those used in Beck (1997). These points are not quite correct. In our experiment (Clahsen et al., 2004), regulars and irregulars were properly balanced, and the account originally proposed by Pinker (1999, pp. 130ff) is meant to explain the production of inflected word forms, and not the results of Beck's particular experiment. Moreover, the relatively fast production latencies for unblocked regulars are, in our view, a consequence of an unimpeded rule route rather than an indication of the rule route applying before blocking (as implied by Indefrey). In production, lexical look-up and the rule route are turned on in parallel. If lexical look-up leads to an existing entry, lemma and lexeme information needs to be extracted (as pointed out by McKee et al.), and this takes time compared to cases for which there are no lexical entries. Blocking is enforced irrespective of timing considerations, by a general principle ensuring that in cases of conflict, specific information (e.g., lexical entries) takes

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

precedence over general rules (compare, e.g., the so-called "elsewhere condition" in linguistics). Thus, in the case of irregulars, successful lexical look-up will block the rule route. In the case of high-frequency regulars, retrieval of a stored entry also blocks the rule route, and this indirectly slows down the rule route relative to low-frequency regulars for which lexical look-up does not require the retrieval of any stored (lemma or lexeme) information.

McKee et al. suggest that adult production models need to be "developmentalized." They propose a developmental account of our findings arguing that the reversed frequency effect is due to incomplete lemmas (for high-frequency regulars) stored in lexical memory yielding a decision conflict with the rule route. Although we agree with their general plea to consider developmental aspects, it should be noted that reversed frequency effects for regulars do not disappear in adults, and that these effects were seen in two age groups of children (5- to 7-year-olds, 11- to 12-year-olds) without there being any clear developmental difference. These findings suggest to us that reversed frequency effects are not specific to the developing production system. Instead, high-frequency regulars (but not low-frequency ones) seem to have memory representations in both children and adults, and in production these will cause a slowdown of the rule route.

Steinhauer discusses the noun–plural data reported from Lück et al. (2001), and claims that the children's relatively low accuracy scores for nouns requiring *-s* plurals do not fit in with the interpretation we gave for the frontal negativities elicited by *-s* plural overregularizations in these children. The former seems to suggest that the *-s* plural rule is not operative in these children, whereas the latter finding was taken to reflect rule-based processing. Note, however, that these experiments were performed with 6- to 12-year-old children and that many previous acquisition studies have shown that correct *-s* plurals are produced much earlier in German child language, that is, from age 1 year, 10 months onward (see Clahsen, Rothweiler, Woest, & Marcus, 1992), and that *-s* plurals are overregularized in child speech and applied under default circumstances from age 3 onward (see Bartke, 1998; Clahsen, Marcus, Bartke, & Wiese, 1996). It is therefore unlikely that the relatively low correctness scores in Lück et al.'s production task reflect any lack of *morphological* knowledge for *-s* plural formation. Instead, these scores are more likely due to the fact that most of the items in this condition were loan words that the children were less familiar with than the native German words we used to create *-s* plural overregularizations in the ERP experiment. Steinhauer expects children who have acquired the *-s* plural rule to produce overregularizations for unfamiliar words in more than 50% of their responses. Note, however, that overregularizations are relatively rare events even in elicitation tasks with novel words. Indeed, the most common response from children in plural elicitation studies using nonce words are unmarked forms, that is, responses in which the children repeat the singular form presented to them without any plural marking. In Berko's (1958) original study, this was the case in approximately 35% of the children's responses, and corresponding studies on German plurals came up with even higher figures (38% in Bartke, 1998; 39% in Schöler & Kany, 1989; and 64% in Gawlitzek-Maiwald, 1994).

In the two sentence-processing experiments we performed with children, differences in the children's WM capacity were found to affect the way temporary

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

ambiguities were resolved and filler-gap dependencies were processed. Traxler correctly points out that WM capacity is a fuzzy concept that covaries with a number of other variables (e.g., lexical decoding skill) and that the contribution of these other variables needs to be established. We did this for the speeded production task reported in Clahsen et al. (2004), in which instead of performing a general WM test, we examined how one specific variable, namely, speed of lexical access, affects children's production latencies of inflected word forms. Given the concerns raised by Traxler, this may turn out to be a more promising approach than directly linking language-processing measures to WM tests.

Overall, we were surprised to see that our commentators largely agreed with our claims about grammatical processing in children. After all, the conclusions we reached (that the child's processing system is basically the same as the adult one and that in on-line comprehension, children rely less on nonsyntactic information than adults) are in contrast to the commonly held belief that semantics comes first in language development. In language processing, children have been argued to rely on perceptual strategies or operating principles (originally proposed by Bever, 1970, and Slobin, 1973, and further developed in much subsequent work) that focus on semantic information and surface properties of the input and allow for direct form–function mappings without invoking any abstract syntactic categories (compare also Bates et al., 1984). It might be the case that these claims apply to children below the age of 5 or 6, that is, younger than those we have studied (see, e.g., Bever, 1970, p. 305). Even though we cannot rule out this possibility, it would produce a rather complicated picture of the development of language processing, with certain points during development at which the system would have to be restructured. Children between the ages of 2 and 5 would not use much grammatical information for language processing but would rely primarily on perceptual strategies and semantic cues. As our results indicate, children from age 6 onward rely more on syntactic information than adults and less on lexical–semantic cues, which would imply a complete turnaround of the earlier system. Later, children would have to restructure the system again in order to include lexical–semantic and pragmatic cues during sentence processing, along with grammatical information. Although it is possible that development of language processing takes a zigzag course of this kind, the alternative idea that the child's processing system is basically the same as the adult one and does not change over time seems to us more straightforward (see also Crain & Wexler, 1999; Fodor, 1998a, 1999; Pinker, 1984, among others).

EVIDENCE FOR SHALLOW STRUCTURES IN L2 PROCESSING

Regarding nonnative language processing, our central claim was that L2 learners who have learned their L2 after acquiring their native language process the L2 differently from native speakers. The SSH claims that during L2 processing, learners compute grammatical representations that lack complex hierarchical structure and abstract, configurationally determined elements such as movement traces, and that nativelike grammatical processing is restricted to "local" domains such as word segmentation or morphosyntactic agreement between closely adjacent constituents. Evidence for shallow processing in the L2 has been found in studies

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

examining syntactic ambiguity resolution and learners' processing of filler-gap dependencies. Learners' sensitivity to argument structure, thematic and plausibility information during L2 sentence processing, in contrast, does not seem to differ much from native speakers'.

Many commentators agreed that the SSH provides an interesting and plausible account of the observed L1/L2 differences in processing. Sekerina and Brooks provide additional evidence from both L1 and L2 studies, suggesting that shallow processing may be more widespread than previously thought, and Sorace thinks that the SSH may also help explain other findings from the L2 literature such as learners' failure to activate the verb phrase internal focus position in L2 Italian. Several commentators (including Birdsong; Gillon Dowens & Carreiras; and Steinhauer) observe that further research is necessary to investigate whether shallow processing also extends to other populations such as early bilinguals, near-native speakers, balanced bilinguals, or L2-dominant speakers. Libben suggests that to probe the possible limits of L2 morphological processing, learners' processing of complex compounds may be worth investigating. Carroll observes that shallow processing should also prevent learners from comprehending subtle semantic distinctions such as scope ambiguities in a nativelike fashion, and Avrutin wonders whether L2 learners might process D-linked *wh*- phrases differently from non-D-linked ones. Gillon Dowens and Carreiras, Libben, and Sabourin note that assessing the possible influence of factors such as proficiency, age of acquisition, individual WM differences, and L1 transfer on L2 processing requires additional, systematic investigation.

Other commentators were more skeptical about our claims. Indefrey asks whether L2 learners do not simply behave like *some* native speakers, notably those with a relatively low WM span. Frenck-Mestre questions some of our argumentation and, like Fernández, raises some methodological criticism of our L2 studies. Steinhauer argues that the SSH is difficult to reconcile with apparent L1 transfer effects in parsing and with the evidence of nativelike performance that has been found in some studies. We address these specific points of criticism first, before turning to more general issues raised by other commentators in the next section.

Given Roberts, Marinis, Felser, and Clahsen's (2006) finding that only native speakers with a relatively high WM capacity showed antecedent reactivation effects in cross-modal priming, Indefrey wonders whether the L2 learners in Marinis, Roberts, Felser, and Clahsen's (2005) study may simply have behaved like low-WM native speakers. The possibility that L2 learners might pattern with a particular WM subgroup of native speakers has been examined in a recent study by Felser and Roberts (2005) on the processing of filler-gap dependencies. Using the same cross-modal picture priming task as did Roberts et al. (2006), Felser and Roberts found that Greek-speaking learners of L2 English behaved differently from both high-WM and low-WM native speakers in that they showed evidence of maintained antecedent activation but not of structurally determined reactivation. That is, the learners showed shorter reaction times to identical than to unrelated targets at both test points, rather than a position-specific antecedent priming effect as was observed in the high-span native speakers. The learners' performance in this task was not influenced by individual WM or proficiency differences, either.

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

These results provide further support for our claim that late learners fail to postulate syntactic gaps when processing unbounded dependencies in their L2 (see also Love, Maas, & Swinney, 2003).

Some of the comments provided by Frenck-Mestre call for clarification. First, it should be noted that contrary to what Frenck-Mestre states, our claim that L2 learners underuse structural information has nothing to do with the observation that N400 effects are often delayed in L2 processing. As the N400 is believed to index lexical–semantic processing, the fact that L2 learners consistently show N400 responses (albeit delayed) is fully in line with the SSH. Instead, ERP evidence supporting the SSH includes the absence of early, left-lateralized anterior negativities in studies of syntactic processing in natural languages, and the observation that P600 responses to syntactic violations are often delayed or absent (compare also Ullman's commentary). We also argued that the nativelike ERP responses observed in Hahne, Müller, and Clahsen's (2006) study on morphological processing indicate that structural processing may be available to adult learners in some domains, such as simple concatenative morphology. Second, Frenck-Mestre draws attention to the fact that existing studies on L2 ambiguity resolution have not always produced consistent results with respect to transfer. Although this is correct, we think that the effects observed by Papadopoulou and Clahsen (2003) provide strong evidence against transfer and input-driven accounts of L2 processing. Highly proficient speakers of L2 Greek who had spent a long time immersed in their L2 still failed to show nativelike ambiguity resolution preferences, despite their native languages showing the same preferences as their L2. This finding is unexpected from the point of view of input-driven/transfer models of L2 acquisition and processing, but not from the perspective of the SSH. For the parser to be able to make any structurally based attachment decisions, sufficiently detailed, hierarchical representations must be available in the first place. Semantically based association, on the other hand, presupposes the presence of relevant lexical cues to interpretation, which we have argued explains the (nativelike) low attachment (NP2) disambiguation preference consistently observed for relative clause (RC) antecedents linked by thematic prepositions.

Fernández also discusses the RC attachment ambiguity, the phenomenon that has been most extensively studied with bilinguals thus far. We do not agree, however, with her assumption that the high attachment preferences observed in many languages are exclusively determined by extrasyntactic factors, notably prosody (including "silent" prosody) and language-specific application of discourse principles. In our view, Gibson, Pearlmutter, Canseco-Gonzalez, and Hickock's (1996) recency/predicate proximity model provides an at least equally plausible, and empirically supported, account of cross-linguistic variation in this domain. In fact, the results from Felser, Roberts, Gross, and Marinis's (2003) and Papadopoulou and Clahsen's (2003) studies are difficult to explain from a prosody/discourse perspective. Note that in both these studies, the segmentation imposed (with a break between the complex antecedent NP and the RC) and the fact that the RCs were relatively long should have biased participants toward high attachment (compare Fodor 1998a; Watson & Gibson, 2002). Despite these potential prosodic biases in our materials, however, our learners did not show any NP1 attachment preferences in either study, and the English native speakers in Felser et al.'s (2003) study

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

actually showed an NP2 preference, in accordance with the recency principle. Nor did our learners seem to transfer the ambiguity resolution preferences from their L1s, as might have been expected, according to Fernández, if they lacked any relevant L2-specific prosodic and/or discourse knowledge.

In addition, we disagree with Fernández' suggestion that off-line results should be more informative than on-line results for the study of RC attachment, and with her dismissing graded acceptability judgments (as used by Papadopoulou & Clahsen, 2003) as a possible method for establishing disambiguation preferences. Acceptability ratings for grammatical sentences that only differ in the way they are disambiguated (NP1 vs. NP2 attachment) are unlikely to reflect anything other than participants' interpretation preferences. Although off-line data provide an indication of ultimately preferred interpretations, on-line methods allow us to determine initial attachment decisions, which may differ from participants' final interpretations (see, e.g., De Vincenzi & Job, 1993). Using on-line methods is essential for understanding what information sources are available to the parser and how comprehenders analyze the input in real time. In our view, the fact that not all on-line studies on RC ambiguities have yielded identical results should be seen as an incentive to further study rather than lead to an exclusive reliance on off-line data (compare also Frenck-Mestre, 2005).

Fernández further observes that knowing an L2 might affect the processing of the first, rendering comparisons between bilinguals and monolinguals problematic (a similar point is made by Carroll). Although it is, of course, conceivable that knowledge of a late-learned L2 affects performance in the L1 (Cook, 2003), given that over half the world's population is estimated to be bilingual to some degree, Fernández' objection seems to call into question the validity of a large number of existing results from L1 processing studies. Does the fact that truly monolingual speakers of, for example, Dutch or Catalan, are virtually nonexistent make it impossible to study ambiguity resolution preferences, or indeed any aspect of language processing, in these languages? We think not. Our main research question in the above studies was whether adult L2 learners process ambiguous sentences in the same way as native speakers of the target language, and the observed L1/L2 differences call for an explanation. Whether learners also differ from monolingual speakers in the way they process their L1 is a different empirical question.

Finally, Fernández' comments on Felser et al.'s (2003) questionnaire results from German- and Greek-speaking learners of L2 English require some clarification. Her observation that the learners showed a majority of NP2 responses overall is, of course, correct, but this is evidently due to their strong NP2 preference for NPs linked by the preposition *with*, which they share with native speakers. What distinguishes the learners from the native speakers is the fact that both learner groups responded at chance level in the *of* condition. Fernández correctly points out that chance-level performance does not necessarily equate lack of a preference. However, if our materials were intrinsically biased toward either NP1 or NP2 attachment, a bias that we took considerable care to avoid, then any such bias should have affected learners and native speakers in a similar way. Yet, only the native speakers showed a reliable NP2 preference in the *of* condition, confirming previous findings for L1 English. Recall that the learners (again, unlike the native speakers) also failed to show any attachment preferences for *of* sentences in our

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

on-line experiments, that is, they behaved consistently across different experimental tasks. Although a direct statistical comparison (as suggested by Fernández) was not possible in Felser et al.'s study due to slight differences between the materials used for each group, Roberts (2003) reports significant differences between Greek-speaking learners' and native English speakers' responses to *of* sentences in her questionnaire task, and a significant interaction of Attachment × Group in the corresponding on-line data, confirming that the learners performed reliably differently from native speakers on ambiguous sentences lacking lexical cues to disambiguation in both tasks.

Steinhauer draws attention to the evidence of nativelike performance and L1 transfer reported in some L2 processing studies (e.g., Sabourin, 2003; Tokowicz & MacWhinney, 2005), which appears to be in conflict with the SSH. In an ERP study with German-, Romance-, and English-speaking learners of Dutch, Sabourin found that learners whose L1 has a similar gender system as the L2 (i.e., the German group) showed a nativelike P600 response to gender violations in Dutch.[1] As the German-speaking learners were also the only ones who had demonstrated above-chance sensitivity to Dutch gender concord in a judgment task, however, it is impossible to tell whether the nativelike P600 observed in the German group was due to L1 influence or a reflection of their relatively higher proficiency in Dutch. Although Steinhauer correctly points out that morphosyntactic agreement involves phrase structure representations, it should be noted that gender concord within the NP is still a very local phenomenon. For subject–verb agreement violations, on the other hand, none of Sabourin's three learners groups showed any P600 effects at all.[2] Note also that, contrary to the Dutch native speakers, the learners showed no early negativities (thought to index automatic structure-building processes) for any of Sabourin's experimental conditions.

P600 effects were also observed by Tokowicz and MacWhinney (2005) in low-proficient, English-speaking learners of L2 Spanish for constructions that are formed similarly in the L1 and the L2, and for those that are unique to the L2. The ERP data revealed that the learners were implicitly sensitive to tense omissions ("similar") and determiner gender violations ("unique") in the L2 but not to determiner number violations ("different" in L1 and L2). The learners' end of sentence judgment accuracy, however, was close to chance for all ungrammatical conditions. The authors conclude that the availability of implicit processing in the L2 depends on the similarity between learners' L1 and the L2. We think that this conclusion is premature, for the following reasons. First, notice that in the absence of native Spanish speakers' control data, it is impossible to tell to what extent the learners' ERP patterns resembled those of monolingual native speakers. Second, the authors do not report any early negativities in their L2 data either, which would have been a stronger indicator of automatic syntactic processing than a P600. Third, exactly why the determiner number condition should fall into the "different" category is not clear. Given that some determiners (indefinites and demonstratives) show overt number agreement in English, English native speakers should be sensitive to determiner number agreement during processing. According to Tokowicz and MacWhinney's own predictions, we would then expect L1 English/L2 Spanish learners to be implicitly sensitive to determiner number violations in the L2, contrary to what they found.[3] Tokowicz and MacWhinney's

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

findings thus do not provide any particularly convincing evidence for nativelike processing or processing transfer.

Steinhauer further wonders whether the results from Hahne et al.'s (2006) study cannot be interpreted in terms of L1 transfer. Hahne et al. found that Russian-speaking learners of German showed a left anterior negativity (LAN)/P600 pattern for overregularizations of past participles but not of noun plurals. However, because past participle formation in Russian is similar to German only in that it involves the same affixes but not in other respects, Hahne et al.'s findings can hardly be considered evidence for L1 transfer. The system of participle formation in Russian differs from the German one in that in Russian, the choice between the three endings is determined by conjugation class and by phonological segments at the right edge of verb stems. In German, on the other hand, the *-t* suffix serves as an overall default, whereas *-n* participle formation only applies to the subclass of strong verbs. In sum, the evidence for L1 transfer in morphosyntactic processing remains sparse, and evidence for nativelike processing in this domain seems to be largely restricted to local mismatches.

ELABORATING THE SSH

A number of comments have signaled a need for some aspects of the SSH to be clarified and elaborated in more detail. Between them, several commentators (including Carroll, Gillon Dowens, & Carreiras, Indefrey, Sekerina, & Brooks, Sorace, Traxler, & Ullman) raise the following questions:

- Under what circumstances does shallow processing occur?
- Is shallow processing restricted to particular linguistic domains?
- Does shallow processing also apply in language production?
- Can shallow processing involve transfer from the L1?
- Are L2 learners restricted to shallow processing, and if so, why?

As Libben points out, another aspect that requires clarification is the question of how exactly the SSH differs from Ullman's (2001) and Paradis' (2004) neurophysiological models of L2 representation and processing. Libben and Steinhauer further ask whether our observations that learners are more nativelike in the processing of inflectional morphology than in syntactic processing may be due to the comparatively simpler materials in Hahne et al.'s (2006) study of morphological processing, rather than being indicative of a more fundamental morphology–syntax dichotomy. Given that shallow processing does not appear to be restricted to L2 learners, Sabourin moreover wonders whether adult learners can really be said to behave in a *qualitatively* different way from native speakers. In the following, we will try to further elaborate and specify the idea of shallow processing in the L2. Naturally, though, we will have to make some assumptions about parsing and the grammar–parser relationship that some may find controversial.

Shallow parsing is a concept familiar from computational approaches to language processing. It is typically thought to involve identifying parts of speech, segmenting the input string into meaningful chunks, and determining what relations these chunks bear to the main verb (compare, e.g., Hammerton, Osborne,

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

Armstrong, & Daelemans, 2002). Evidence for shallow parsing in the L1 (e.g., Christianson, Hollingworth, Halliwell, & Ferreira, 2001; Ferreira, 2003; Ferreira, Bailey, & Ferraro, 2002; Sanford & Sturt, 2002) is compatible with processing models that assume that comprehension normally involves both the application of semantically based comprehension heuristics and full syntactic analyses. According to the integrated processing model proposed by Townsend and Bever (2001), for example, the L1 comprehension mechanism normally assigns two different kinds of representations to an input string, a rough and ready "pseudosyntax" representation based on lexical information and statistical patterns, and a fully specified syntactic description. Whereas the former allows comprehenders to quickly determine a sentence's likely meaning, the latter serves to supplement and confirm the analysis. Although there are several aspects of Townsend and Bever's model that we consider problematic, its basic tenet that native speakers "understand sentences twice" may provide a useful template for understanding L2 processing.[4] Let us assume, then, that the human language processing system makes available two different routes for computing sentence interpretations, which usually work in parallel. The full parsing route is fed by the grammar (a system of symbolic rules and principles of structure building), but shallow processing is guided by lexical–semantic and pragmatic information, world knowledge, and strong associative meaning or form patterns.

We argued that what distinguishes nonnative comprehenders from native ones is that in L2 processing, the shallow processing route predominates. Why should this be so? Basically, we can see two possibilities. One possibility is that the same parsing mechanisms that are used in L1 processing (such as minimal attachment, recency, or the active filler strategy) are also available in L2 processing, but that their application is restricted due to the knowledge source that feeds the structural parser, the L2 grammar, being incomplete, divergent, or of a form that makes it unsuitable for parsing. The second possibility is that although the L2 grammar is sufficiently detailed and suitable for parsing, full parsing fails due to the unavailability or deficiency of the required parsing mechanisms.[5] In line with previous suggestions made by Epstein, Flynn, and Martohardjono (1996) and others, Sorace seems inclined toward the second possibility, interpreting the findings reported in our target article to mean that "some of the differences between native and (advanced) nonnative speakers may be at the level of grammatical processing, rather than grammatical representations." In contrast to Sorace, however, we think that the first possibility is more realistic, for the following reasons. First, there are both learnability (e.g., Fodor, 1998a, 1999; Gibson & Wexler, 1994) and empirical reasons (e.g., De Vincenzi & Job, 1993; Frazier, 1987, Gibson et al., 1996; Gibson, Pearlmutter, & Torrens, 1999) for assuming that basic parsing mechanisms are universal, and thus do not have to be learned. If this is correct, then parsing principles such as minimal attachment or the active filler strategy that guide L1 processing should also be available in L2 processing. In contrast, language-specific properties of the L2 grammar must obviously be learned. Second, there is evidence that learners develop interlanguage grammars that are fundamentally different from L1 grammars (e.g., Bley-Vroman, 1990; Clahsen & Muysken, 1986, 1989). In short, we believe that although both processing routes are available to L2 learners in principle, successful structural parsing depends on the availability

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

(and accessibility) of sufficiently detailed, implicit grammatical knowledge. The idea that the full parsing route is underused in L2 processing due to inadequacies of the L2 grammar is illustrated in Figure 1.

With the full parsing route being of limited use in L2 processing, learners' interpretations will typically be derived via the shallow processing route only.[6] The consistent absence of early LAN effects in ERP studies on L2 sentence processing might be taken to suggest that the stage at which initial structures are built automatically on the basis of word category information is skipped altogether in nonnative comprehension. L2 processing may thus be said to differ qualitatively from L1 processing in that native speakers but not L2 learners will normally carry out a full parse as well. Recall that learners' use of lexical and plausibility information in L2 ambiguity resolution is well documented, and several ERP studies have shown that learners' brain responses to lexical–semantic violations are essentially nativelike (see Mueller, 2005, for a review). Adult learners' ability to use metalinguistic information, world knowledge, and pragmatic inferencing, and to match associatively stored meaning and form patterns to the input, will further help them to become generally successful L2 comprehenders. Under this view, whether or not L2 learners can also develop nativelike parsing abilities will depend on their acquiring a nativelike grammar.

Grammatical knowledge also informs language production, and to the extent that production and comprehension make use of the same processing mechanisms, the SSH applies to production also. However, as language production is much more under the speaker's conscious control, effects of shallow processing in production may be more difficult to spot. If learners' reliance on shallow processing ultimately reflects inadequacies of the L2 grammar, we would further expect that individual WM differences, as opposed to, for example, factors like proficiency or amount of exposure, have little or no effect on L2 parsing performance. Although few studies have investigated the influence of WM on L2 processing, the results available thus far seem to confirm this prediction (see Felser & Roberts, 2005; Juffs, 2004, 2005; Sato & Felser, 2005).

Next, let us return to the question of L1 transfer in processing. Given the processing model outline above, we would expect L1 transfer to influence L2 processing only indirectly, as a consequence of one or more of the knowledge sources that feed the processing system being affected by properties of the L1. Frenck-Mestre and Pynte's (1997) observation that ambiguity resolution in the L2 was influenced by argument structure differences between L1/L2 translation equivalents, for instance, provides an example of lexically based transfer (see also Juffs, 1998). Much of the research within the competition model moreover suggests that the degree to which learners exploit different types of surface cue in L2 comprehension may be influenced by properties of their L1 (MacWhinney, 1997). To our knowledge, clear evidence of the L1 competence grammar affecting the real-time parsing of complex grammatical structures in the L2, on the other hand, has not yet been found.

Finally, it should be noted that the SSH differs from Ullman's (2001) and Paradis' (2004) models in several respects. Contrary to these models, the SSH is a psycholinguistic hypothesis that remains essentially neutral with respect to the question of the possible neurophysiological localization of shallow versus deep

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

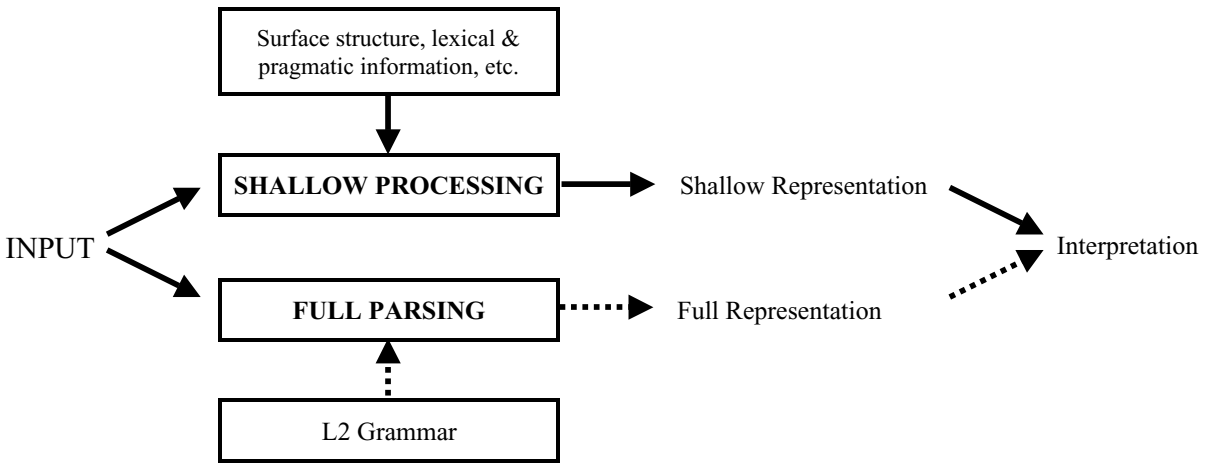


Figure 1. Of the two routes to interpretation available in principle, full parsing is restricted in L2 sentence processing because of inadequacies of the L2 grammar.

processing. We also do not necessarily subscribe to the idea that learners draw predominantly on declarative knowledge sources when processing their L2. Shallow processing may well involve the application of procedural knowledge, such as pragmatic inferencing. The SSH further differs from the above models in that it differentiates between relatively simple morphological rules (which learners may be able to employ in a nativelike fashion) and the computation of complex syntactic representations (which is predicted to remain problematic even for advanced L2 learners). Note, however, that in contrast to what Ullman says, the SSH does not rule out the possibility that for some bilingual populations, nativelike performance may extend to linguistic domains other than those mentioned above, or in our article. We would expect, however, that although "proceduralization" (which, in the present context, might be understood as referring to learners' use of the full parsing route) may be possible for some learners, the availability of full parsing will typically remain restricted even at later stages of L2 learning.

THEORETICAL AND PRACTICAL IMPLICATIONS

Turning to more general issues, Duffield points out that given that the grammar and the parser are closely intertwined, a clear distinction between grammatical competence and parsing performance may not, in fact, be possible, and Juffs wonders how the concept of shallow processing might fit with current theories of grammar such as Chomsky's (1995) minimalist framework. Several commentators have raised questions regarding the implications of the SSH for language development. Carroll and Juffs point out that the SSH should ultimately be integrated into a more comprehensive theory of L2 acquisition or learning. Birdsong, Gillon Dowens, and Carreiras and Libben all ask whether L2 learners can ever acquire nativelike parsing routines, and Gillon Dowens and Carreiras wonder whether this may be subject to a critical period. Juffs, furthermore, asks about the possible pedagogical implications of the SSH.

We would like to reemphasize that contrary to Duffield's understanding of the SSH, we do not claim that the observed L1/L2 performance differences reflect mere processing differences, nor do we assume that "a particular piece of linguistic performance [. . .] is uniquely due to the grammar or to the processing system." As outlined above, we do, in fact, think that the opposite holds true: L2 processing is different because of inadequacies of the L2 grammar. That is, the L2 parser will be unable to successfully apply even universal processing mechanisms (such as minimal attachment) if the L2 grammar fails to provide sufficient grammatical information. Although most of the learner groups we examined had demonstrated a high level of L2 proficiency, their being able to provide nativelike off-line judgments on the structures under investigation does not imply that the nature and extent of their grammatical knowledge was nativelike. On-line tasks are believed to reduce the degree to which participants are able to draw on "explicit" grammatical knowledge during processing, which is why we think it important to supplement off-line data with corresponding on-line data.

In response to Juffs' question of whether L2 learners are capable of feature checking in the sense of Chomsky (1995), we would like to point out that being able to establish a semantic link between, for example, a fronted *wh-* phrase and

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

its subcategorizer during comprehension does not, in our view, imply that any checking of formal (specifically, uninterpretable) features takes place. As successful feature checking is usually thought to depend on properties of configurational structure such as c-command, we would expect that during shallow processing, nonlocal checking of formal (as opposed to semantic) features will not normally be possible.

We agree with Carroll and Juffs that theories of language acquisition are incomplete unless they also incorporate assumptions about processing. Given the shortage of empirical data bearing on this issue, however, current models of how grammatical competence and parsing performance may be linked in development are primarily based on theoretical considerations (see, among others, Crain & Wexler, 1999; Fodor, 1998a, 1998b, 1999; Gibson & Wexler, 1994; Truscott & Sharwood Smith, 2004). An in-depth discussion of these models is beyond the purpose and scope of this response, however. Note that for child L1 acquisition, the continuity of parsing hypothesis that we argued for is consistent with the "parsing to learn" approach to grammar building advocated by Fodor (1998a, 1999) and others, which provides a solution to the acquisition paradox mentioned in our article. For adult L2 learners, on the other hand, acquisition through parsing will be a much more limited option if, as we have argued, L2 learners predominantly use the shallow processing route to interpretation.

The extent to which some learner groups may nevertheless achieve nativelike parsing performance remains to be determined. Gillon Dowens and Carreiras cite evidence suggesting that highly proficient bilinguals who have spent a long time immersed in the L2 may process both gender and number agreement in a similar way to native speakers (Gillon-Dowens, Barber, Vergara, & Carreiras, 2004), and Birdsong mentions results from Golato (2002) that indicate that some L2-dominant bilinguals may use nativelike word segmentation strategies. Observe, however, that as in Hahne et al.'s (2006) and Sabourin's (2003) studies, the domains in which nonnative speakers show evidence of nativelike processing are again local ones, in the sense specified above.

Some studies have shown that age of acquisition is a crucial factor in L2 processing (e.g., Weber-Fox & Neville, 1996), suggesting that the availability of the full parsing route in L2 acquisition may be subject to a critical period. Whether the reduced availability of full parsing in late L2 acquisition is ultimately due to neurobiological changes occurring around puberty (as has been suggested by Ullman, 2004), or whether this is a secondary consequence of the increase in size and accessibility of relevant extragrammatical knowledge during adolescence, we are unable to tell.

Finally, let us briefly consider the possible implications of the SSH for language teaching, a point raised by Juffs. If the SSH is correct, then it does indeed seem that a stronger than usual focus on formal properties of the L2 grammar (rather than on pragmatics) is called for. To the extent that "processing instruction" (VanPatten, 1996, 2004) or other "focus on form" techniques (e.g., Long & Robinson, 1998; Williams, 1995) can help learners develop a more nativelike L2 grammar, they will also pave the way for nativelike processing performance. Observe, however, that the above conclusion holds true only if the attainment of nativelike, implicit competence and processing abilities are considered to be important goals of L2

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

learning. As several studies have shown, many L2 learners manage to develop virtually nativelike comprehension (and possibly also, production) abilities without necessarily showing nativelike processing performance. In the absence of any comparative studies investigating the effect of different teaching methods on learners' parsing abilities, it is still unclear to what extent full parsing can be taught. Depending on the definition of learning goals, time limitations, and other constraints that formal language instruction may be subject to, some teachers may wish to prioritize on vocabulary building, comprehension strategies, and communicative skills rather than on grammar or processing instruction. Ostriches may have lost their ability to fly, but they can be pretty good runners.

As a concluding note, we would like to highlight the *comparative* approach of our research program, which we think turned out to be extremely useful. Different linguistic phenomena (morphology, syntax) were studied in two groups of language learners (children and adults) using a variety of experimental methods. An approach that relies on different experimental methods is likely to avoid, or at least reduce, uncertainties arising from weaknesses of individual techniques, gaps in particular data sets, or potentially confounding factors. In our view, the comparative investigation of child L1 and adult L2 processing is particularly useful because, if child L1 processing largely involves full parsing whereas much of L2 grammatical processing is "shallow," then comparing L1 and L2 learners' processing performance will allow us to systematically study the properties of the two processing routes, for example, how much of a language can be processed by shallow parsing? Precisely which constructions require a full parse? Finally, the comparison of morphological and syntactic phenomena in processing has revealed interesting differences and similarities. Had we restricted our investigations to either morphological or syntactic processing, our conclusions would have been different, at least for L2 learners. Clearly, however, the studies we reported only represent the beginnings of an emerging field of research, and further comparative studies of different learner groups, different languages, and different linguistic phenomena will be necessary to achieve a better understanding of grammatical processing in language learners.

NOTES

1. Foucart and Frenck-Mestre (2004) report that local gender mismatches also elicited a P600 effect in proficient German-speaking learners of French.
2. Further evidence for L2 learners' lack of sensitivity to subject–verb agreement violations during processing includes the results from a reading-time study by Jiang (2004) with Chinese-speaking learners of English.
3. Note further that some asymmetries in Tokowicz and MacWhinney's (2005) materials make it difficult to compare the three ungrammatical conditions directly. Their tense omission condition involved sentences that lacked a finite auxiliary, that is, sentences that were incomplete. This was not the case in the two other conditions, both of which involved a local feature mismatch. The gender agreement condition moreover differed from the other two in that the critical word was in sentence-final position, which raises the possibility that the P600 effect here reflects end of sentence wrap up processes.

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

4. Contrary to Townsend and Bever (2001), we do not assume, for example, that the initial semantic analysis (or shallow processing, in our terms) normally precedes full parsing. Townsend and Bever's model, moreover, differs from ours in the level of syntactic detail attributed to their "pseudosyntax" representations, the computation of which, according to the authors, also involves "movement of *wh*-argument gaps [*sic*] into their source location" (p. 228). The argument dependency model proposed by Schlesewsky and Bornkessel (2004) also incorporates two parallel processing routes, a "thematic" and a "syntactic" one. Their model differs from both Townsend and Bever's model and the one outlined here in several respects, though. A detailed critique of these models is beyond the scope of this reply, however.
5. The issue of the grammar–parser relationship is still far from settled. Existing proposals range from the idea that the internalized competence grammar *is* the parser (Phillips, 1996, 2003; Weinberg, 1999) to the claim that "grammar" does not exist except as a mere epiphenomenon reflecting the workings of a statistical parser (Bates & MacWhinney, 1989). We have adopted the fairly standard view here that the grammar feeds the parser, and that parsing is guided by additional, "least-effort" based processing principles (compare, e.g., Crain & Fodor, 1985). Note that unlike the grammar, parsing is subject to time constraints and capacity limitations. Computationally complex sentences such as center-embedded structures, for example, tend to be difficult to process even though they are licensed by the grammar.
6. Observe further that L2 learners' overreliance on shallow processing is unlikely to be conductive to the development of implicit L2 knowledge (or "proceduralization," in Ullman's terms). As Gillon Dowens and Carreiras put it, shallow processing could be "an early interlanguage feature of L2 sentence processing that continues to be effective, and so employed, even at advanced learner stages." It should be noted that, contrary to what Gillon Dowens and Carreiras state, we do not assume that the learners we tested were necessarily steady or end-state learners. Rather, we examined advanced learners who had demonstrated nativelike or near-native knowledge of the relevant grammatical domains in off-line tasks. As mentioned earlier, the extent to which end-state learners, learners at the top end of the proficiency scale or L2-dominant learners exhibit nativelike processing behavior remains yet to be established.

REFERENCES

Bartke, S. (1998). *Experimentelle Studien zur Flexion und Wortbildung*. Tübingen: Niemeyer.

Bates, E., & MacWhinney, B. (1989). Functionalism and the competition model. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing* (pp. 3–73). New York: Cambridge University Press.

Bates, E., MacWhinney, B., Caselli, C., Devescovi, A., Natale, F., & Venza, V. (1984). A cross-linguistic study of the development of sentence interpretation strategies. *Child Development, 55*, 341–354.

Beck, M.-L. (1997). Regular verbs, past tense and frequency: Tracking down a potential source of NS/NNS competence differences. *Second Language Research, 13*, 93–115.

Berko, J. (1958). The child's learning of English morphology. *Word, 14*, 150–177.

Bever, T. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), *Cognition and the development of language* (pp. 279–352). New York: Wiley.

Bley-Vroman, R. (1990). The logical problem of second language learning. *Linguistic Analysis, 20*, 3–49.

Chomsky, N. (1995). *The minimalist program*. Cambridge, MA: MIT Press.

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

Christianson, K., Hollingworth, A., Halliwell, J., & Ferreira, F. (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42*, 368–407.
Clahsen, H., Hadler, M., & Weyerts, H. (2004). Speeded production of inflected words in children and adults. *Journal of Child Language, 31*, 683–712.
Clahsen, H., Marcus, G., Bartke, S., & Wiese, R. (1996). Compounding and inflection in German child language. In G. Booij & J. van Marle (Eds.), *Yearbook of morphology 1995* (pp. 115–142). Dordrecht: Kluwer.
Clahsen, H., & Muysken, P. (1986). The accessibility of universal grammar to adult and child learners: A study of the acquisition of German word order. *Second Language Research, 2*, 93–119.
Clahsen, H., & Muysken, P. (1989). The UG paradox in L2 acquisition. *Second Language Research, 5*, 1–29.
Clahsen, H., Rothweiler, M., Woest, A., & Marcus, G. (1992). Regular and irregular inflection in the acquisition of German noun plurals. *Cognition, 45*, 225–255.
Cook, V. (2003). *Effects of the second language on the first*. Clevedon: Multilingual Matters.
Crain, S., & Fodor, J. D. (1985). How can grammars help parsers? In D. Dowty, L. Karttunen, & A. Zwicky (Eds.), *Natural language parsing: Psychological, computational and theoretical perspectives* (pp. 94–128). Cambridge: Cambridge University Press.
Crain, S., & Wexler, K. (1999). Methodology in the study of language acquisition: A modular approach. In W. Ritchie & T. Bhatia (Eds.). *Handbook of child language acquisition* (pp. 387–425). San Diego, CA: Academic Press.
De Vincenzi, M., & Job, R. (1993). Some observations on the universality of the late-closure strategy. *Journal of Psycholinguistic Research, 22*, 189–206.
Epstein, S., Flynn, S., & Martohardjono, G. (1996). Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences 19*, 677–714.
Felser, C., & Roberts, L. (2005). *Processing wh-dependencies in a second language: A cross-modal priming study*. Manuscript submitted for publication.
Felser, C., Roberts, L., Gross, R., & Marinis, T. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics, 24*, 453–489.
Ferreira, F. (2003). The misinterpretation of non-canonical sentences. *Cognitive Psychology, 47*, 164–203.
Ferreira, F., Bailey, K., & Ferraro, V. (2002). Good enough representations in language comprehension. *Current Directions in Psychological Science, 11*, 11–15.
Fodor, J. D. (1998a). Learning to parse? *Journal of Psycholinguistic Research, 27*, 285–319.
Fodor, J. D. (1998b). Parsing to learn. *Journal of Psycholinguistic Research, 27*, 339–374.
Fodor, J. D. (1999). Triggers for parsing with. In E. Klein & G. Martohardjano (Eds.), *The development of second language grammars: A generative approach* (pp. 373–406). Amsterdam: John Benjamins.
Foucart, A., & Frenck-Mestre, C. (2004, September). *Processing of grammatical gender information in French as first and second language*. Poster presented at AMLaP, Aix-en-Provence.
Frazier, L. (1987). Syntactic processing: evidence from Dutch. *Natural Language and Linguistic Theory, 5*, 519–559.
Frenck-Mestre, C. (2005). Eye-movement recording as tool for studying syntactic processing in a second language: A review of methodologies and experimental findings. *Second Language Research, 21*, 175–198.
Frenck-Mestre, C., & Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native languages. *Quarterly Journal of Experimental Psychology, 50A*, 119–148.
Gawlitzek-Maiwald, I. (1994). How do children cope with variation in the input? The case of German plural and compounding. In R. Tracy & E. Lattey (Eds.), *How tolerant is universal grammar? Essays on language learnability and language variation* (pp. 225–266). Tübingen: Niemeyer.
Gibson, E., Pearlmutter, N., Canseco-Gonzalez, E., & Hickock, G. (1996). Recency preferences in the human sentence processing mechanism. *Cognition, 59*, 23–59.
Gibson, E., Pearlmutter, N., & Torrens, V. (1999). Recency and lexical preferences in Spanish. *Memory & Cognition, 27*, 603–611.
Gibson, E., & Wexler, K. (1994). Triggers. *Linguistic Inquiry, 25*, 407–454.
Gillon-Dowens, M., Barber, H., Vergara, M., & Carreiras, M. (2004, September). *Does practice make perfect? An ERP study of morphosyntactic processing in highly proficient English–Spanish late bilinguals*. Poster presented at AMLaP, Aix-en-Provence.

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

Golato, P. (2002). Word parsing by late-learning French–English bilinguals. *Applied Psycholinguistics, 23*, 417–446.
Hahne, A., Müller, J., & Clahsen, H. (2006). Morphological processing in a second language: Behavioral and ERP evidence for storage and decomposition. *Journal of Cognitive Neuroscience, 18*.
Hammerton, J., Osborne, M., Armstrong, S., & Daelemans, W. (2002). Introduction to special issue on machine learning: Approaches to shallow parsing. *Journal of Machine Learning Research, 2*, 551–558.
Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics, 25*, 603–634.
Juffs, A. (1998). Some effects of first language argument structure and syntax on second language processing. *Second Language Research, 14*, 406–424.
Juffs, A. (2001). Psycholinguistically-oriented L2 research. In M. McGroarty (Ed.), *Annual review of applied linguistics*. Cambridge: Cambridge University Press.
Juffs, A. (2004). Representation, processing, and working memory in a second language. *Transactions of the Philological Society, 102*, 199–225.
Juffs, A. (2005). The influence of first language on the processing of *wh-* movement in English as a second language. *Second Language Research, 21*, 121–151.
Long, M., & Robinson, P. (1998). Focus on form. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 15–41). Cambridge: Cambridge University Press.
Love, T., Maas, E., & Swinney, D. (2003). The influence of language exposure on lexical and syntactic language processing. *Experimental Psychology, 50*, 204–216.
Lück, M., Hahne, A., Friederici, A., & Clahsen, H. (2001, November). *Developing brain potentials in children: An ERP study of German noun plurals*. Paper presented at 26th Boston University Conference on Language Development.
MacWhinney, B. (1997). Second language acquisition and the competition model. In A. De Groot & J. Kroll (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp. 113–142). Mahwah, NJ: Erlbaum.
Marinis, T., Roberts, L., Felser, C., & Clahsen, H. (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition, 27*, 53–78.
Mueller, J. (2005). Electrophysiological correlates of second language processing. *Second Language Research, 21*, 152–174.
Papadopoulou, D., & Clahsen, H. (2003). Parsing strategies in L1 and L2 sentence processing: A study of relative clause attachment in Greek. *Studies in Second Language Acquisition, 24*, 501–528.
Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.
Phillips, C. (1996). *Order and structure*. Unpublished PhD dissertation, MIT.
Phillips, C. (2003). Linear order and constituency. *Linguistic Inquiry, 34*, 37–90.
Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
Pinker, S. (1999). *Words and rules*. New York: Basic Books.
Roberts, L. (2003). *Syntactic processing in learners of English*. Unpublished PhD dissertation, University of Essex, Colchester.
Roberts, L., Marinis, T., Felser, C., & Clahsen, H. (2006). Antecedent priming at gap positions in children's sentence processing. *Journal of Psycholinguistic Research, 35*.
Sabourin, L. (2003). *Grammatical gender and second language processing: An ERP study*. Unpublished PhD dissertation, University of Groningen.
Sanford, A., & Sturt, P. (2002). Depth of processing in language comprehension: Not noticing the evidence. *Trends in Cognitive Science, 6*, 382–386.
Sato, M., & Felser, C. (2005, May). *Sensitivity to different types of information in L2 sentence processing: Evidence from speeded grammaticality judgements*. Presented at the Annual Meeting of the Japan Second Language Association.
Schlesewsky, M., & Bornkessel, I. (2004). On incremental interpretation: Degrees of meaning accessed during sentence comprehension. *Lingua, 114*, 1213–1234.
Schöler, H., & Kany, W. (1989). Lernprozesse beim Erwerb von Flexionsmorphemen: Ein Vergleich sprachbehinderter mit sprachunauffälligen Kindern am Beispiel der Pluralmarkierung (Untersuchung I und II). In G. Kegel (Ed.), *Sprechwissenschaft & Psycholinguistik 3. Beiträge aus Forschung und Praxis* (pp. 123–176). Opladen: Westdeutscher Verlag.

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press

Slobin, D. (1973). Cognitive prerequisites for the development of grammar. In C. Ferguson & D. Slobin (Eds.), *Studies of child language development* (pp. 175–208). New York: Holt, Rinehart & Winston.

Tokowicz, N., & MacWhinney, B. (2005). Implicit and explicit measures of sensitivity to violations in second language grammar: An event-related potential investigation. *Studies in Second Language Acquisition, 27*, 173–204.

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

Truscott, J., & Sharwood Smith, M. (2004). Acquisition by processing: A modular perspective on language development. *Bilingualism: Language and Cognition, 7*, 1–20.

Ullman, M. (2001). The neural basis of lexicon and grammar in first and second language: The declarative/procedural model. *Bilingualism: Language and Cognition, 4*, 105–122.

Ullman, M. (2004). Contributions of memory circuits to language: The declarative/procedural model. *Cognition, 92*, 231–270.

VanPatten, B. (1996). *Input processing and grammar instruction*. Chestnut Hill, NJ: Ablex.

VanPatten, B. (2004). Input processing in second language acquisition. In B. VanPatten (Ed.), *Processing instruction: Theory, research, and commentary* (pp. 5–31). Mahwah, NJ: Erlbaum.

Watson, D., & Gibson, E. (2002, March). *When does prosody influence parsing?* Poster presented at the 15th Annual CUNY Conference on Human Sentence Processing, New York.

Weber-Fox, C., & Neville, H. (2001). Sensitive periods differentiate processing of open- and closed-class words: An ERP study of bilinguals. *Journal of Speech, Language and Hearing Research, 44*, 1338–1353.

Weinberg, A. (1999). A minimalist theory of human sentence processing. In S. Epstein & N. Hornstein (Eds.), *Working minimalism* (pp. 283–315). Cambridge, MA: MIT Press.

Williams, J. (1995). Focus on form in communicative language teaching: Research findings and the classroom teacher. *TESOL Journal, 7*, 6–11.

https://doi.org/10.1017/S0142716406060206 Published online by Cambridge University Press