Harald Clahsen | Loay Balkhair | John-Sebastian Schutter
Ian Cunnings

The time course of morphological processing in a second language

Suggested citation referring to the original publication:
Second Language Research 1 (2012)
DOI http://doi.org/10.1177/0267658312464970
ISSN (online) 1477-0326
ISSN (print) 0267-6583

Article

second
language
research

Second Language Research
29(1) 7–31

DOI: 10.1177/0267658312464970
slr.sagepub.com
SAGE

The time course of morphological processing in a second language

Harald Clahsen
University of Potsdam, Germany

Loay Balkhair
University of Essex, UK

John-Sebastian Schutter and Ian Cunnings
University of Edinburgh, UK

Abstract

We report findings from psycholinguistic experiments investigating the detailed timing of processing morphologically complex words by proficient adult second (L2) language learners of English in comparison to adult native (L1) speakers of English. The first study employed the masked priming technique to investigate *-ed* forms with a group of advanced Arabic-speaking learners of English. The results replicate previously found L1/L2 differences in morphological priming, even though in the present experiment an extra temporal delay was offered after the presentation of the prime words.

The second study examined the timing of constraints against inflected forms inside derived words in English using the eye-movement monitoring technique and an additional acceptability judgment task with highly advanced Dutch L2 learners of English in comparison to adult L1 English controls. Whilst offline the L2 learners performed native-like, the eye-movement data showed that their online processing was not affected by the morphological constraint against regular plurals inside derived words in the same way as in native speakers. Taken together, these findings indicate that L2 learners are not just slower than native speakers in processing morphologically complex words, but that the L2 comprehension system employs real-time grammatical analysis (in this case, morphological information) less than the L1 system.

Keywords

compounds, derivational morphology, English as a seond language, inflectional morphology, late bilinguals, masked priming, morphology processing, past tense, shallow structure hypothesis

---

Corresponding author:
Harald Clahsen, Potsdam Research Institute for Multilingualism, University of Potsdam, Karl-Liebknecht-Strasse 24–25, 14476 Potsdam, Germany
Email: harald.clahsen@uni-potsdam.de

1 Introduction

Much current psycholinguistic research focuses on the temporal dynamics of language processing by specifying which sources of information the language processing system employs at particular moments in time. One controversial issue in this research concerns the question of to what extent language processing is serial in nature – relying, for example, on form-level information earlier than on meaning – and to what extent it is parallel, in that the system employs different information sources at the same time; see e.g. Jackendoff (2007) for discussion. Although most of this research has been done on sentence-level phenomena, specifically ambiguity resolution (e.g. Lewis, 2000), time-course issues also arise for morphological processing.

One case in point concerns processes involved in the recognition of morphologically complex words. Consider, for example, morphological priming experiments with first language (L1) speakers (Marslen-Wilson, 2007). On the one hand, overt priming experiments in which both primes and targets are presented long enough to be consciously recognized produced robust priming effects for semantically transparent pairs (e.g. *calmness–calm*) but not for semantically opaque pairs (e.g. *department–depart*). Masked priming experiments, however, in which the prime is only shown very briefly – too short to be consciously recognized – revealed priming effects for both transparent and opaque word pairs. To account for these apparently conflicting results, masked and overt priming have been claimed to be sensitive to different stages of processing. Masked priming appears to tap an early pre-lexical stage of word recognition at which a prime word's form and morphological structure is accessible but not its semantic properties; hence the same (masked) priming effect for *department–depart* as for *calmness–calm*. Overt priming, by contrast, includes a later stage of lexical processing at which the different semantic properties of these word pairs are recognized; hence the reduced (overt) priming effect for semantically opaque prime–target pairs.

Another case of how the timing of different information sources is important for morphological processing concerns complex word forms that combine inflectional, derivational, and compounding processes. Combinations between these processes are constrained in several ways. Affixes may, for example, have specific selectional restrictions (Fabb, 1988; Plag, 2002), such as denominal *-y* suffixation that does not apply to suffixed words (**duckling-y*). There are also general morphological, semantic and (possibly) phonological constraints on the occurrence of inflected forms (Anderson, 1992; Kiparsky, 1982; Stump, 2001; Wiese, 1996). The avoidance of regular but not irregular plurals inside compounds is the most well-known example of this kind; see Gordon (1985) and much subsequent work. Regular plurals are not allowed inside compounds while irregular plurals are marginally acceptable and singular non-heads are preferred (e.g. *rats eater, mice eater, rat/mouse eater*). This asymmetry has been attributed to three constraints. The dispreference for plural forms, irrespective of regularity, has been argued to result from a constraint against compound modifiers with plural number semantics (Haskell et al., 2003), while the avoidance of regular plurals has been explained either in terms of a morphological constraint against regularly inflected, grammatically-computed, compound modifiers (Berent and Pinker, 2007), or a phonological one against non-heads with sibilant-final codas (Seidenberg et al., 2007). Evidence from eye-movement monitoring during reading

(Cunnings and Clahsen, 2007) indicates that these constraints become operative at different points in time during compound processing, with, for example, the morphological constraint affecting earlier eye-movement measures than the semantic constraint.

Although non-native second language (L2) processing has recently received much attention (e.g. Clahsen and Felser, 2006a; Gor, 2010; Perani, 2005; VanPatten and Jegerski, 2010), the precise time course of when L2 learners access different types of information sources during processing is still largely unknown. With respect to this question, two main proposals are currently discussed. One hypothesis claims that native and non-native processing employ the same mechanisms but that L2 processing may require extra time and is generally slower (e.g. McDonald, 2006). Others have proposed more substantial differences between native and non-native processing arguing, for example, that L2 processing relies more on lexical memory and less on the procedural system than L1 processing (Ullman, 2005) and that adult L2 learners' ability to make use of grammatically-based parsing strategies is reduced relative to their sensitivity to lexical-semantic and other non-structural information cues (e.g. Clahsen and Felser, 2006b). Against this background, the present article aims to shed light on the time course of L2 morphological processing by presenting new findings from behavioural and eye-movement experiments. We examined the role of processing speed and of how different kinds of linguistic cues affect morphological processing over time. Two morphological phenomena of English were studied: regularly inflected *-ed* forms in Experiments 1 and 1a, and derived suffixed forms containing plural versus singular base nouns in Experiments 2 and 3.

II Recognizing inflected word forms during L2 processing

The first experiment examined early automatic processes involved in the recognition of regularly inflected *-ed* forms using the masked visual priming technique. In a masked priming experiment, a prime word is briefly presented immediately followed by a target word or non-word, and participants have to decide whether the target is an existing word or a non-word. The short prime presentation times do not usually allow participants to consciously recognize the prime. For native speakers, several studies using this technique found morphological priming effects for inflected and derived word forms in different languages that could be dissociated from any facilitation due to the orthographic and/or semantic overlap between primes and targets (Rastle et al., 2000, 2004).

There are a few studies that have employed the masked priming technique to compare L1 and L2 processing. For inflectional morphology, most previous experiments revealed non-native-like masked priming patterns for L2 learners. Silva and Clahsen (2008) found morphological priming for *-ed* forms in L1 English, but not in German, Chinese, and Japanese L2 learners of English, a finding that was interpreted as indicating that L2 learners rely less on combinatorial morphological processing than L1 speakers. Likewise, Neubauer and Clahsen (2009) obtained significant priming effects for regular *-t* participles in L1 German, but not in Polish L2 learners of German. On the other hand, Feldman et al. (2009) reported results from a masked priming experiment in which a subgroup of the L2 learners they tested showed a priming effect for regularly inflected primes. Note, however, that the corresponding L1 English control data did not show any significant prime-by-condition interactions, a surprising result given the robust morphological priming effects that have been reported in several other masked priming experiments for L1 speakers. This raises the question of whether the pattern Feldman et al. (2009) obtained for L2 learners is indeed equivalent to L1 morphological priming. Leaving aside this concern, it is true that across studies masked priming effects for inflected word forms appear to be less robust in L2 than in L1 processing. While this finding may be indicative of substantial L1/L2 differences, as suggested by Silva and Clahsen (2008), another possibility would be that the results are due to slowed processing in the L2 than the L1. Specifically, L2 learners may require more time than native speakers for processing of the prime word. Hence, the design chosen in previous masked priming experiments in which the target was presented immediately after the prime word may not have provided enough time for L2 learners to morphologically parse the prime word. In the present study, we examined groups of advanced Arabic-speaking learners of English on *-ed* forms using the masked priming technique. There were two experimental versions – a 'standard' version modelled after Silva and Clahsen (2008) and a 'delayed' version that included an extra temporal delay after the presentation of the prime words – to compensate for L2 learners' overall slower processing speed. In a follow-up experiment, potential priming effects resulting from orthographic and semantic relatedness between primes and targets were tested.

1 Participants

In the main experiment (Experiment 1), we compared data from the 21 native speakers of English (18 females) reported by Silva and Clahsen (2008), to a new group of 20 L2 learners of English with Arabic as L1 (3 females), 10 of whom underwent the standard and 10 the delayed version of the masked priming paradigm. The follow-up experiment (Experiment 1a) tested a new group of 21 native speakers of English (mean age: 23.6, 14 women), none of whom participated Experiment 1. Appendix 1 presents more detailed information about the L2 participants. They had all been living in the UK at the time of testing, but did not consider themselves as balanced bilinguals. They achieved proficiency scores in the grammar part of the Oxford Placement Test (Allan, 1992) which put them into the 'advanced' or 'upper advanced' bands (mean: 79.1, range: 75–90). Whilst two participants had a score of 90, the remaining ones scored between 75 and 80. All L1 and L2 participants had normal or corrected-to-normal vision, were never diagnosed with any learning or other behavioural disorder, and were naive with respect to the purpose of the experiments.

2 Materials

The materials for Experiment 1 were taken from Silva and Clahsen's (2008) Experiments 1 and 2. There were 21 critical prime–target pairs in the morphologically 'Related' condition in which the target word (PRAY) was preceded by its corresponding *-ed* (*prayed*) form, 21 prime–target pairs in the 'Unrelated' condition, in which there was no morphological, orthographic, or semantic relation between the prime and target word (*wash*–PRAY), and 21 pairs in the 'Identity' condition (*pray*–PRAY). The targets in all conditions were the unmarked bare stems, which were presented in upper case letters, unlike the primes presented in lower case, to minimize visual overlap between primes and targets. All target words were monosyllabic, four letters long and had a mean stem/word frequency of 42.4 (per million) in the CELEX database (Baayen et al., 1993). The primes were also closely matched for frequency and length. The critical prime–target pairs were distributed over three experimental lists that were matched as closely as possible in terms of mean word form and stem frequencies. Each list contained 21 critical prime–target pairs, 7 targets that were preceded by an Identity, 7 by a Related and 7 by an Unrelated prime with each target appearing only once in each version. In addition to the critical prime–target pairs, 303 filler trials were included into each of the three experimental lists; see Silva and Clahsen (2008: 249).

The purpose of the follow-up Experiment 1a was to directly compare morphological priming effects with potential priming effects resulting from the semantic and/or formal overlap between primes and targets. For the morphologically related condition, Experiment 1a used the same prime–target pairs as Experiment 1, e.g. *prayed*–PRAY. The semantic control condition consisted of 21 semantically but not orthographically or phonologically related prime–target pairs taken from a synonymy dictionary (Sinclair, 1993). The 21 prime–target pairs were selected from an initial set of 42 pairs, on the basis of a pretest in which 13 native speakers of English rated word pairs on a five-point scale ranging from ‘almost identical in meaning’ (5) to ‘have nothing in common in their meaning’ (1). All 21 pairs included in the priming experiment (see Appendix 2) received a high semantic similarity score of 4 or above. The formal control condition examined the role of orthographic relatedness between primes and targets. There were 21 orthographically but not semantically or morphologically related prime–target pairs (see Appendix 3). As in the morphologically related condition, the target words were fully contained in the primes, but unlike in the Related condition, the additional letters in the prime words of the formal overlap condition did not represent English affixes, e.g. *yellow–yell*. The mean formal overlap, i.e. the average proportion of concatenative letters in the prime that also appeared in the target and vice versa (Rastle et al., 2000: 512) was .71 (SD: .07), which is not significantly different from the mean formal overlap in the morphologically related condition (.69, SD: .075). The critical prime–target pairs were distributed over different experimental versions, so that each target appeared only once in each version. In addition to the critical prime–target pairs, a set of 275 filler pairs was included, which consisted of 56 word–word pairs (of which 28 were morphologically related other than through *-ed* forms), 57 non-word–word pairs, 81 word–non-word pairs, and 81 non-word–non-word pairs. Half of the non-word filler primes were orthographically related to their targets. The stimulus set of each version consisted of 324 prime–target pairs, 85.5% of them were unrelated. The prime–target pairs were pseudo-randomized in the same way as in the main experiment.

3 Procedure, data scoring, analysis

Two different masked priming designs were compared in Experiment 1. The first design is familiar from most previous studies of this kind (Forster and Davis, 1984) and involves three visual events. First, a forward mask consisting of a series of Xs appeared on the screen for 500 ms (which also served as a fixation point) followed by the prime (displayed for 60 ms), which was then immediately followed by the presentation of the target item for 500 ms. The second design introduced a delay of 200 ms between primes and targets, as illustrated in Figure 1.

As in the standard design, a forward mask consisting of a series of Xs appeared on the screen for 500 ms followed by the prime (displayed for 60 ms). Unlike the standard design, however, a blank screen was shown at the offset of the prime for 200 ms followed by the target word. Whilst maintaining a masked priming design (in which the prime was still only presented for 60 ms), this modification allowed us to examine whether extra time given for the processing of the prime word had any effects on processing.[1]

In Experiments 1 and 1a, primes and targets were presented in size 36 font in white letters against a black background on a 16 inch (406 mm) screen. In order to reduce the amount of pure visual priming, the prime words were presented in lower and the target words in upper case letters. Participants were not informed of the presence of any prime stimuli. The experiments began with a short practice session consisting of 10 prime–target pairs.

To determine whether participants were aware of the prime words, participants were asked what they saw on the screen, after they performed the experiments. In the experimental versions that used the standard design (in which the target word immediately followed the prime), most participants reported seeing the screen flash or flicker at times, and a few participants reported seeing sequences of letters before the target. Some participants reported seeing a word between the mask and the target word (5 out of the 31 in Experiment 1 and 11 out of the 21 in Experiment 1a), but none of them was able to correctly identify any of the primes. In the delayed design, however, in which the target was presented 200 ms after the offset of the prime, all participants reported seeing some printed material before the targets. This could be due to the sharp contrast between the white font of the prime words against the black background shown for 200 ms, which may have produced a reflection of the prime word.

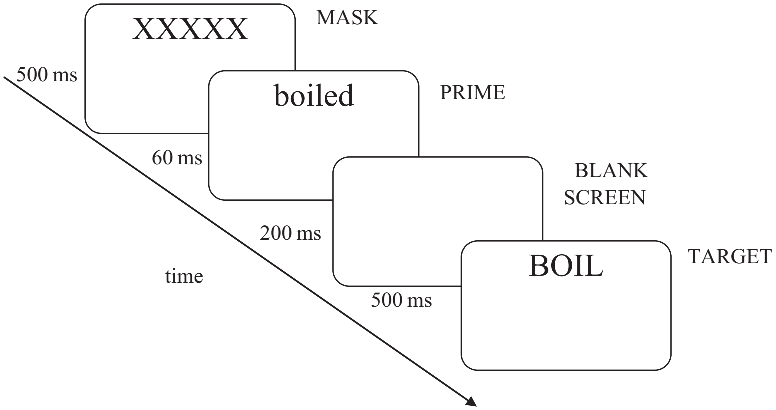

Figure 1 The delayed masked priming experiment

The experiments were performed in a dedicated, quiet room and lasted for approximately 60 to 70 minutes for the L2 groups and 25 to 35 minutes for the L1 groups. The L1 groups were allowed to take one break during the experiment, while the L2 groups were allowed to take two breaks. Before the experiment, participants were given detailed instructions about the task. After the experiment, all L2 participants took a vocabulary test to ensure that they knew the meanings of the critical items. In this test, participants were asked to select from four choices the correct synonym for the critical target words. There were no mistakes in this vocabulary test confirming that the L2 participants were familiar with the meanings of the critical items. Prior to the calculation of lexical decision times, incorrect responses, i.e. non-word responses to existing words and word responses to non-words, and outliers, i.e. extreme reaction times exceeding two SDs from a participant's mean per condition, were excluded from any further analysis. The reaction time (RT) and error data were submitted to analyses of variance (ANOVA) followed by planned comparisons if appropriate. Whilst mean RTs and log-transformed RTs are shown, statistical analyses were performed on the latter. The *p*-values of all analysis were Greenhouse–Geisser corrected for non-sphericity whenever applicable.

4 Results

Experiment 1 presents data from a group of 20 L2 learners, 10 of whom were tested with the standard design in which targets immediately followed primes ('L2-I') and 10 with the (200 ms) delayed design ('L2-II'). These data were compared to the group of 21 L1 English speakers ('L1-control') reported by Silva and Clahsen (2008), who were tested with the standard priming design. Table 1 displays mean RTs to the targets (as well as standard deviations and error rates) in the three conditions of the masked priming experiment. Planned comparisons are shown in Table 2.

We first compared the L2 learners taken as one group to the L1 controls. ANOVAs with the factors Group (L2, L1) and Prime Type (Identity, Related, Unrelated) on the error data revealed a significant main effect of Group ($F_1(1,39) = 22.06$, $p < .001$; $F_2(1,40) = 24.37$, $p < .001$), due to the L2 learners' higher error rates. There were no

Table 1 Raw and log-transformed mean RTs (in ms), SDs (in parentheses), and error rates

		L2	L2-I	L2-II	L1*
Identity	RTs	674 (209)	768 (201)	573 (197)	451 (42.6)
	Log RTs	6.46 (0.35)	6.64 (0.16)	6.33 (0.27)	6.11 (0.09)
	Errors	10.7%	10%	11.4%	1.4%
Related	RTs	757 (227)	818 (204)	697 (284)	463 (52.8)
	Log RTs	6.58 (0.33)	6.69 (0.16)	6.51 (0.27)	6.11 (0.18)
	Errors	10.7%	12.9%	8.6%	2.7%
Unrelated	RTs	790 (232)	834 (181)	747 (260)	518 (80.8)
	Log RTs	6.63 (0.3)	6.7 (0.16)	6.58 (0.25)	6.28 (0.15)
	Errors	14.3%	12.9%	15.7%	6.8%

Note: * from Silva and Clahsen (2008: 251)

Table 2 Planned comparisons for the L2 versus L1 groups in Experiment 1

	L2	L1
Identity-Unrelated	$t_1(19) = 5.35, p < .001$	$t_1(20) = 4.7, p < .001$
	$t_2(20) = 4.06, p < .001$	$t_2(20) = 4.62, p < .001$
Related-Identity	$t_1(19) = 3.74, p < .001$	$t_1(20) < 1$
	$t_2(20) = 2.2, p < .05$	$t_2(20) < 1$
Related-Unrelated	$t_1(19) = 1.31, p = .20$	$t_1(20) = 6.9, p < .001$
	$t_2(20) = 1.3, p = .21$	$t_2(20) = 4.01, p < .001$

further main effects or interactions for either participants or items (all *p*s > .1). For the RT data, the ANOVAs yielded main effects of Prime Type ($F_1(2,78) = 30.6, p < .001$; $F_2(2,80) = 17.03, p < .001$) and Group ($F_1(1,39) = 57.8, p < .001$; $F_2(1,40) = 261.7, p < .001$), and, more importantly, an interaction of Group by Prime Type, that was reliable by participants and marginal by items ($F_1(2,78) = 5.4, p < .01$; $F_2(2,80) = 2.41, p = .096$). Further examination of this interaction (see Table 2) showed repetition priming effects for both participant groups, i.e. significantly shorter target RTs after an identical than an unrelated prime. By contrast, the patterns for morphologically related prime–target pairs were different in the L1 and the L2 groups. In the L2 group, the Related and Unrelated conditions produced similar (target) RTs, both of which were significantly longer than in the Identity condition. This suggests that morphologically related primes did not produce any priming in the L2 groups. This differs from the L1 group (Silva and Clahsen, 2008), where the Related condition yielded similar (target) RTs as Identity primes, which were significantly shorter than those for the Unrelated prime condition, indicating that morphologically related primes produced a full priming effect.

To examine potential differences between the two designs of the masked priming experiment, we also compared the data from the two subgroups of L2 learners. ANOVAs with the factors Group (L2-I, L2-II) and Prime Type (Identity, Related, Unrelated) on the error data did not reveal any main effects or interactions (all *Fs* < 1). The ANOVAs for the RT data yielded main effects of Prime Type ($F_1(2,36) = 18.7, p < .001$; $F_2(2,78) = 8.26, p < .001$) and Group ($F_1(1,18) = 4.52, p < .05$; $F_2(1,39) = 36.03, p < .001$), and an interaction of Group by Prime Type for participants but not for items ($F_1(2,36) = 6.7, p < .01$; $F_2(2,78) = 1.94, p = .15$). The source of this interaction is between-group differences in the magnitudes of the repetition priming effect. Although both groups showed a significant repetition priming effect (see Table 3), the magnitude of repetition priming, i.e. the log RT difference between the Unrelated and the Identity conditions, was significantly larger in the L2-II than the L2-I group ($t_1(18) = 3.67, p < .01$; $t_2(39) = 1.77, p = .09$). With respect to morphologically related prime–target pairs, the results were parallel for both groups. There was no significant morphological priming and similar magnitudes of facilitation (0.01 versus 0.07: $t_1(18) < 1$; $t_2(29) < 1$). These results are incompatible with the view that L2 processing is slower but otherwise parallel to L1 processing. Instead, the comparison between the two priming designs shows that extra time given to

Table 3 Planned comparisons for the two L2 subgroups in Experiment 1

	L2-I	L2-II
Identity-Unrelated	$t_1(9) = 3.06$, $p < .05$	$t_1(9) = 7.28$, $p < .001$
	$t_2(19) = 2.27$, $p < .05$	$t_2(20) = 3.67$, $p < .05$
Related-Identity	$t_1(9) = 2.35$, $p < .05$	$t_1(9) = 3.53$, $p < .05$
	$t_2(19) = 1.69$, $p = .11$	$t_2(20) = 2.04$, $p = .06$
Related-Unrelated	$t_1(9) < 1.$	$t_1(9) = 1.22$, $p = .25$
	$t_2(20) < 1$	$t_2(20) = 1.22$, $p = .24$

Table 4 Mean RTs (in ms), SDs (in parentheses), and error rates in Experiment 1a

		MORPH	SEM	FORM
Related	RTs	525 (96)	571 (109)	600 (139)
	Errors	0%	4.7%	4.6%
Unrelated	RTs	575 (114)	542 (94)	594 (122)
	Errors	5%	3.9%	7%

process a morphologically complex word does not yield (masked) morphological priming effects in L2 processing.

Another question that needs to be addressed is whether the priming effect for pairs such as *prayed*–PRAY reported in Experiment 1 for the L1 group is indeed morphological in nature. To this end, Experiment 1a tested a new group of L1 speakers on semantic and formal overlap control conditions (‘SEM’, ‘FORM’), in addition to the same morphological test condition as in Experiment 1 (see Table 4). ANOVAs with the factors Condition (MORPH, SEM, FORM) and Prime Type (Related, Unrelated) on the RT data yielded a main effect of Condition ($F_1(2,40) = 11.96$, $p < .001$; $F_2(2,56) = 10.69$, $p < .001$) and a significant interaction between Prime Type and Condition ($F_1(2,40) = 12.63$, $p < .001$; $F_2(2,56) = 6.26$, $p < .01$).[2] This interaction was followed up by planned comparisons (see Table 5 below), which showed a significant priming effect for the morphological condition ($t_1(20) = 4.52$, $p < .001$; $t_2(20) = 3.18$, $p < .01$), and no reliable differences for either the semantic ($t_1(20) = 1.65$, $p = .11$; $t_2(18) = 1.79$, $p = .09$) or the formal overlap condition ($t_1(20) = 1.4$, $p = .18$; $t_2(18) < 1$). These results indicate that the priming effect in the MORPH condition is indeed due to the morphological relatedness between primes and targets and cannot be attributed to their semantic or orthographic overlap.

The priming results can be summarized as follows. First, the L2 group exhibited significant repetition priming effects in the absence of any reliable morphological priming for *-ed* forms. Second, the L2 learners did not produce a morphological priming effect, either in the immediate or the delayed design that provided extra time for processing a morphologically complex prime word. These findings replicate previously reported L1/L2 differences from masked priming studies of inflectional morphology (Neubauer and Clahsen, 2009; Silva and Clahsen, 2008). In addition, the present set of results suggests that these differences are not due to slowed L2 processing.

Table 5 Mean ratings, SDs (in parentheses) for derived word forms in Experiment 2

		L2	L1-controls*
IRREGULAR	singular	5.10 (1.26)	5.15 (1.22)
	plural	4.60 (1.49)	4.12 (1.35)
REGULAR	singular	5.24 (1.46)	5.07 (1.12)
	plural	3.46 (1.50)	2.87 (1.23)
Phonological condition	similar	5.15 (1.37)	5.08 (1.02)
	dissimilar	5.12 (1.31)	5.07 (1.09)

Note: * from Cunnings and Clahsen (2008: 158)

III The timing of morphological constraints in L2 processing

The second set of studies (Experiments 2 and 3) examines how constraints against inflected forms inside derived words affect L2 processing. The specific linguistic phenomenon under study was derived words of English containing singular versus plural base nouns. Results from acceptability judgments with L1 speakers (Cunnings and Clahsen, 2008) showed that derived words containing singular base nouns (*ratless*) are preferred over those with plural ones and that amongst the latter, derived words with regular plurals as base nouns (*ratsless*) were rated significantly worse than those with irregular plurals (*liceless*). These contrasts can be explained in terms of two constraints (Cunnings and Clahsen, 2008). The first one attributes the preference for singular noun bases of derived words to a Category Constraint that restricts the kinds of morphological types that may enter derivational processes to stems. Consequently, singular forms being identical to uninflected stems in English are preferred inside derived word forms, whereas all kinds of inflected words are dispreferred. The second constraint is concerned with the dislike of regular plurals inside compounds and derived words. This can be attributed to a (morphological) Structure Constraint – originally proposed by Kiparsky (1982) – according to which regular inflection should not feed either compounding or derivation; see also Di Sciullo and Williams (1987), Wiese (1996). These two constraints in tandem account for the pattern of acceptability ratings obtained for L1 English native speakers. Derived words containing uninflected stems (*flealess*) do not violate any of these constraints and are fully acceptable, whilst those containing regular plurals violate both constraints and are ungrammatical (**fleasless*), and derived words containing irregular plurals violate the Category but not the Structure Constraint and as such are marginally acceptable (?*liceless*). As an alternative to the Structure Constraint, Haskell et al. (2003) proposed that the phonological properties of regular plural nouns are responsible for their ban inside compounds, specifically a dislike of modifiers with sibilant-final codas. For L1 English speakers, several studies have disconfirmed this idea, both as an account of the plurals-in-compounds effect (Berent and Pinker, 2007) and for the dispreference of regular plurals in derived words (Cunnings and Clahsen, 2008).

Experiments 2 and 3 examined whether highly advanced Dutch L2 learners of English are sensitive to these constraints. Offline performance was studied with an acceptability judgment task (Experiment 2), and online processing by measuring eye movements during reading (Experiment 3), a technique that provides a rich source of data on moment-to-moment language processing (Rayner, 1998).

Both experiments were administered to the same group of 25 Dutch L2ers (17 females); see Appendix 1. The control data come from Cunnings and Clahsen (2008), 40 L1 speakers of English (29 females) for Experiment 2 and 24 (12 females) for Experiment 3. Although the L2 learners had a relatively early age of onset of English (mean: 9;6, range: 4;0–15;0), they did not consider themselves as balanced bilinguals. They had all been living in the UK at the time of testing and achieved high proficiency scores (mean: 88.9, range: 69–98) in the grammar part of the Oxford Placement Test. Eighteen participants scored 88 and above, which puts them into the 'very advanced' to 'expert user' bands, 5 participants were in the 'advanced' (81–85), and 2 in the 'upper intermediate' range (69, 73). For all participants, Experiment 3 preceded Experiment 2. All L1 and L2 participants had normal or corrected-to-normal vision and no reported learning or other behavioural disorders. They were naive with respect to the purpose of the experiments.

1 Acceptability judgements

The procedures and materials of this experiment were taken from Cunnings and Clahsen (2008) who adopted the acceptability rating task from previous studies of compounding (e.g. Berent and Pinker, 2007; Cunnings and Clahsen, 2007; Haskell et al., 2003) to examine derived words.

a Materials and procedure. The experimental items were derived words presented in short two-sentence contexts. The derived words consisted of the adjective-forming affix *-less* and the adverb-forming affix *-wise* paired with a nominal stem. The critical set of experimental items tested the acceptability of derived words containing stems with singular or plural forms of base nouns that take regular and irregular inflections, e.g. *louse–lice* versus *flea–fleas*. The critical set of experimental items contained 72 derived words and four types of closely matched nominal stems (irregular singular/plural, regular singular/plural). Examples of how stimuli were presented to participants are given in (1):

(1) a. Following recent food scares regarding British meat, the Government has advised farmers to slaughter much of their livestock. One local farmer has been left completely **OX(EN)/PIG(S)-LESS** since the cull was recommended.
 b. Many dog owners have problems keeping their pets free from mites and other itchy pests. **LOUSE(LICE)/FLEA(S)-WISE**, our dog has luckily never had such problems.

In addition, there was a set of 32 items for a 'phonological condition' to test for whether surface-form properties affect acceptability judgments (Haskell et al., 2003). These items all had singular forms as nominal stems and were affixed with *-less* or *-wise*. Half of the nominal stems ended in /s/, and the other half had other codas, e.g. *fox* versus *wolf*. The items of the phonological condition were also embedded in sentences similar to those in (1). Items in both sets were controlled for frequency, length in letters and number of syllables; see Appendices 4 and 5 taken from Cunnings and Clahsen (2008: 174f.).

Four pseudo-randomized lists of 34 sentence contexts (18 experimental, 16 control ones) and 40 filler sentences were presented to participants for acceptability judgment, so that no experimental items from the same condition appeared adjacent to each other. Participants were asked to rate the items shown in bold on a scale ranging from 1 (highly unacceptable) to 7 (highly acceptable).

b Results. Table 5 presents mean ratings and standard deviations from the L2 group in comparison to the group of L1 English speakers from Cunnings and Clahsen (2008). As regards the regular/irregular conditions, three-way ANOVAs for the factors Regularity (irregular/regular), Number (singular/plural) and Group (L2/L1) revealed significant main effects of Regularity ($F_1(1,63) = 37.83, p < .001$; $F_2(1,136) = 14.93, p < .001$) and Number ($F_1(1,63) = 84.75, p < .001$; $F_2(1,136) = 88.90, p < .001$) in both the participant and item analyses, and for Group ($F_1(1,63) = 1.30, p = .259$; $F_2(1,136) = 4.56, p < .05$) in the item analysis only, as well as an interaction of Regularity and Number ($F_1(1,63) = 33.31, p < .001$; $F_2(1,136) = 17.44, p < .001$). There were no further main effects or interactions. Planned comparisons indicate that (1) regular plurals were significantly more acceptable inside derived words than both irregular plurals ($t_1(64) = 6.99, p < .001$; $t_2(34) = 5.77, p < .001$) and singulars ($t_1(64) = 10.10, p < .001$; $t_2(34) = 9.61, p < .001$) and that (2) irregular plurals were less acceptable than corresponding singulars ($t_1(64) = 5.32, p < .001$; $t_2(34) = 3.63, p < .005$). For the phonological condition, an additional ANOVA with the variables Similarity (similar, dissimilar) and Group (L2/L1) did not reveal any significant main effects or interactions (all $Fs < 1$). Taken together, these results indicate similar judgment patterns for both participant groups, higher ratings for derived word forms containing singular than plural forms, with particularly low ratings for derived words containing regular plurals, and no difference for the phonological condition, irrespective of whether the nominal stems had codas that were phonologically similar (e.g. *fox*) or dissimilar (e.g. *wolf*) to those of regular plural forms.

Although the above analyses did not produce interactions with Group, the data in Table 5 indicate that the contrast in acceptability ratings between singular and plural forms is considerably weaker in the L2 than in the L1 group, due to plurals receiving relatively high ratings. Thus it could be that the significant contrasts from the above-reported pairwise comparisons between regular and irregular plurals and between plurals and singulars are mainly carried by the L1 group. To address this concern, separate ANOVAs and planned comparisons (Table 6) were performed on the L2 data on their own.

These analyses revealed the same significant interaction of Regularity and Number ($F_1(1,24) = 13.42, p < .001$; $F_2(1,68) = 9.7, p < .005$) and main effects (Regularity: $F_1(1,24) = 14.194, p < .005$; $F_2(1,68) = 4.05, p < .05$; Number: $F_1(1,24) = 16.76, p < .001$), $F_2(1,68) = 26.06, p < .001$) that Cunnings and Clahsen (2008: 158) reported for the L1 data. These results together with the planned comparisons in Table 6 confirm that both participant groups prefer singular stems inside derived forms over any kind of plural form. In addition, both participant groups judged derived forms with irregular plural stems as significantly better than those with regular plural as stems.

Table 6 Planned comparisons of the acceptability judgments in Experiment 2

Regular plural × Singular		Regular plural × Irregular plural		Irregular plural × Singular	
$t_1(64) =$	$t_2(34) =$	$t_1(64) =$	$t_2(34) =$	$t_1(64) =$	$t_2(64) =$
10.10**	9.61**	6.99**	5.77**	5.32**	3.63*

Notes: ** $p < .01$; * $p < .05$; $^{(*)}$ $p < .1$

These results show that highly proficient Dutch L2 learners judged the kinds of derived words tested in native-like ways. Furthermore, our finding that the phonological properties of singular base nouns (in particular sibilant-final codas) had no reliable effect on the L2 learners' acceptability ratings of the derived words is also parallel to what was found for derived words in L1 English (Cunnings and Clahsen, 2008) as well as with Berent and Pinker's (2007) observation that compounds such as *fox chaser* are rated as being fully acceptable by native speakers of English.

These findings suggest that in an offline task, highly proficient Dutch L2 learners of English appear to be sensitive to both the Category Constraint against plural noun bases and the morphological Structure Constraint against regular plurals in the same way as native speakers of English.

2 Eye movements during reading

This experiment examines eye-movements during reading as a window into the time-course of processing of the different kinds of derived words tested in Experiment 2. Cunnings and Clahsen (2008) found that L1 English speakers' preference of uninflected stems inside derived word forms is visible early on during processing, through significantly reduced first fixation and gaze durations for derived words with uninflected (singular) stems compared to those with plural forms as stems. They also found effects of the morphological constraint that bans regular plurals from occurring inside derived word forms, on later eye-movement measures, in that forms such as *fleasless* elicited significantly longer rereading times than corresponding forms with irregular inflected stems (e.g. *liceless*). Against the finding from Experiment 2 that highly proficient Dutch L2 learners of English were sensitive to these constraints offline in the same way as L1 English speakers, the purpose of the present experiment was to find out whether these L2 learners show the same native-like sensitivity to these constraints during online processing.

a Materials, procedure, and data analysis. The materials were taken from Experiment 2 of Cunnings and Clahsen (2008), which compared derived words containing singular, regular plural and irregular plural base nouns. In this experiment we only included one 'singular' condition, for two reasons: first, because there were no differences between the different singular stems in Experiment 2 ('reg-singular', 'irreg-singular') and, second, to ensure that conditions were matched for word length, a lexical property that is known to affect online reading times. Furthermore, because the phonological condition did not produce any effect in Experiment 2, it was not included. This resulted in a 3 × 2 design with the factors 'Condition' ('reg-plural', 'irreg-plural', 'singular') and Group (L1, L2). For each triplet of experimental items (e.g. *oxenless, animalsless* and *horseless*), three different context paragraphs were constructed, within which each critical item could be placed felicitously. The first sentence of these context paragraphs acted as a lead-in providing the context and the third sentence as a wrap-up absorbing potential spill-over effects. An example context paragraph is shown in (2):

(2) Our dog was pest infested until we treated him with a special shampoo.
He's since been completely **LICE/MITES/FLEALESS** and has stopped itching all over.
If he has similar problems in the future we now know what to do.

The experimental item set consisted of 54 novel derived words in three conditions, containing irregular plurals (e.g. *liceless*), regular plurals (*mitesless*), singular nouns (*flealess*). The experimental items were controlled for frequency, length in letters, and number of syllables, and were pre-tested with a group of L1 English speakers to ensure that they displayed the intended range of acceptability; see Appendix 6 (Cunnings and Clahsen, 2008: 175). Sixty filler texts comprised of a variety of different syntactic constructions were added and presented in sentence contexts parallel to those of the experimental items.

The 114 experimental and filler texts were pseudo-randomized and presented to participants in one of three lists such that each participant saw each of the 54 experimental items and each context paragraph only once, and that no texts containing experimental items from the same condition appeared adjacent to each other. Participants were instructed to read the context paragraphs to themselves at their normal reading pace. After approximately 20% of the trials, participants had to answer a content question to check whether they had paid attention to the materials. Both participant groups had high accuracy rates for these content questions, with a mean of 86.2% (SD: 10.85) for the L2 group tested here, indicating that participants paid attention to and understood the contents of the sentences. Eye-movements were recorded with the same apparatus that Cunnings and Clahsen (2008) used for the L1 study, an EYELINK II® eye-tracker (SR Research Ltd., Canada) at 500 Hz. Few data had to be removed before the analysis, 1.56% of the critical trials in the L2 dataset, which included cases of track loss and of fixations of the target derived word region before reading of the second sentence had begun.

Reading time measures were calculated for the derived word region, which included the derived word itself plus half a letter space either side. The same five reading-time measures (first fixation duration, gaze duration, regression path duration, rereading time, total viewing time) and the same procedures that Cunnings and Clahsen (2008) employed for the L1 data were used to analyse the L2 data.

b Results. Table 7 presents mean reading times (and standard deviations) for five eye-movement measures. These measures each index progressing later stages of processing. 'First fixation' durations refer to the duration of the first fixation within a target region and only include the eyes' first encounter of a target region. 'Gaze duration' additionally

Table 7 Mean durations (and standard deviations) of five reading time measures for the derived word region in Experiment 3

	L2			L1*		
	Regular	Irregular	Singular	Regular	Irregular	Singular
First fixation durations	272 (55)	285 (54)	262 (44)	255 (52)	258 (35)	237 (42)
Gaze durations	376 (110)	385 (107)	328 (74)	363 (137)	366 (120)	297 (90)
Regression path durations	433 (116)	466 (144)	425 (107)	467 (186)	423 (124)	389 (114)
Rereading times	226 (117)	171 (139)	121 (76)	224 (168)	126 (81)	120 (81)
Total viewing times	601 (180)	556 (191)	449 (115)	587 (210)	492 (131)	417 (143)

Note: * from Cunnings and Clahsen (2008: 166)

includes all subsequent fixations during the first pass. First fixation and gaze duration both measure the processes involved with and immediately after lexical access. ‘Regression path durations’ include additional processing resulting from regressions back to earlier parts of the text before moving past the target word to the right. In contrast to regression path durations which do not include any rereading of the target region after it has been exited to the right, ‘rereading times’ include fixations of the target region as a result of regressing back from later parts of the sentence, thus reflecting later, second-pass stages of processing. Finally, ‘total viewing times’, the summed duration of all fixations of the target region, are a general index of processing load (see Rayner, 1998).

Statistical analyses – 3 × 2 ANOVAs with the factors Condition (REG, IRR, SG) and Group (L1, L2) – revealed main effects of Condition for all five measures reflecting overall shorter RTs for singular than for plural stems in both groups (first fixation durations: $F_1(2,94) = 12.66, p < .001$; $F_2(2,51) = 7.93, p = < .005$; gaze durations: $F_1(2,94) = 29.83, p < .001$; $F_2(2,51) = 6.06, p < .005$; regression path durations: $F_1(2,94) = 6.57, p < .005$; $F_2(2,51) = 2.48, p = .094$; rereading times: $F_1(2,94) = 20.50, p < .001$; $F_2(2,51) = 10.78, p < .001$; total viewing times: $F_1(2,94) = 41.48, p < .001$; $F_2(2,51) = 10.85, p < .001$). In addition, there was a significant main effect of Group for first fixation durations that was marginal in the participant analysis ($F_1(2,48) = 3.24, p = .078$; $F_2(2,51) = 37.28, p < .001$). There were no further main effects or interactions for any of the five eye-movement measures.

Although there were no interactions with the factor Group, it is important to compare offline and online performance in the L2 groups parallel to Cunnings and Clahsen’s (2008: 167f.) analyses of the L1 data. We therefore performed the same analyses for Experiment 3 as for Experiment 2, examining the L2 data from the current experiment on their own using a series of one-way ANOVAs for the five eye-movement measures with the three-level factor Condition followed by planned comparisons. These analyses yielded significant effects for four of the five measures in the L2 data, first fixation duration ($F_1(2,48) = 5.95, p < .01$; $F_2(2,51) = 5.01, p < .05$), gaze duration ($F_1(2,48) = 13.02, p < .001$; $F_2(2,51) = 2.90$; $p = .064$), rereading time ($F_1(2,48) = 12.97, p < .001$; $F_2(2,51) = 7.41, p < .005$) and total viewing time ($F_1(2,48) = 20.96, p < .001$; $F_2(2,51) = 7.57, p < .005$), but not for regression path duration ($F_1(2,48) = 1.78, p = .179$; $F_2(2,51) = .98, p = .381$). The results of the planned comparisons are shown in Table 8.

Table 8 Planned comparisons for the L2 data on five reading time measures in Experiment 3

	Regular × Singular		Regular × Irregular		Irregular × Singular	
	$t_1(24)$ =	$t_2(70)$ =	$t_1(24)$ =	$t_2(70)$ =	$t_1(24)$ =	$t_2(70)$ =
First fixation	1.54	1.53	2.25	1.66	2.99	3.31
Durations	R = S	R = S	R < I*	R < I(*)	I > S**	I > S*
Gaze	3.47	1.89	.92	.20	4.74	2.89
Durations	R > S**	R > S$^{(*)}$	R = I	R = I	I > S**	I > S**
Rereading	5.38	3.63	2.39	1.91	2.63	2.10
Times	R > S**	R > S*	R > I*	R > I$^{(*)}$	I > S*	I > S*
Total viewing	5.69	3.70	1.89	1.13	5.05	2.93
Times	R > S**	R > S**	R > I$^{(*)}$	R = I	I > S**	I > S**

Notes: ** $p < .01$; * $p < .05$; $^{(*)}$ $p < .1$

The results from Table 8 can be summarized in three points. First, derived words containing irregular plural base nouns attracted significantly longer first fixation durations than those with regular plural ('I > R') and singular base nouns ('I > S'). Second, gaze durations elicited significantly longer RTs for derived words with plural than with singular base nouns ('R/I > S'), but no difference between regular and irregular plural ones ('R = I'). Finally, rereading and total viewing times also yielded significant differences between singular base nouns on the one hand and plural ones on the other ('R/I > S'), but less reliable differences between regular and irregular base nouns ('R(>) = I'). If L2 processing was simply slower than L1 processing, we would expect to find the same reading-time patterns as observed in the L1 data reported by Cunnings and Clahsen (2008), albeit in later reading time measures. This was not the case, however. For native speakers, Cunnings and Clahsen (2008) found that derived words with irregular plural bases did not yield longer reading times than those with regular ones, on any measure. Instead, first fixations and gaze durations in the L1 data revealed longer reading times for derived words with plural bases (irrespective of regularity) than those with singular bases, with no differences between regular and irregular plurals. This pattern of reading times was followed by elevated reading times for derived words with regular plural bases relative to those with singular and irregular bases on measures that include later processes (rereading times, total viewing times). Thus, L1 and L2 processing show different reading-time patterns.

These results indicate both L1/L2 similarities and differences as to how the constraints against plural nouns inside derived words are employed during L1 and L2 processing. In both L1 and L2 reading, gaze durations elicited shorter reading times for derived words with singular base nouns than for those with plural base irrespective of regularity, indicating sensitivity to the Category Constraint. Differences between L1 and L2 processing were found with respect to the morphological Structure Constraint. L1 processing was affected by this constraint in that elevated rereading times were observed for derived words with regular plural bases in comparison to those with irregular plural or singular bases, with no differences between the latter two. In the L2 data, however, this pattern of results was not observed in any of the reading-time measures.

IV General discussion

In the following, we discuss the results from the four experiments reported above focusing on aspects of the timing of processing morphologically complex words in an L2.

1 *Parallel versus serial processing*

Studies investigating the temporal dynamics of language raise the question of whether particular information sources are restricted to particular stages or moments in time during processing or whether the system accesses all relevant information sources in parallel. The results reported in the present study are more in line with the former view and suggest that morphological processing is not entirely parallel, either in the L1 or the L2.

Consider the priming results. Experiment 1 along with other masked priming studies (e.g. Neubauer and Clahsen, 2009; Silva and Clahsen, 2008) did not show any facilitation for regularly inflected prime words in different groups of advanced L2 learners. Experiments using overt priming designs, however, revealed similar amounts of facilitation for *-ed* primes as native speakers (Basnight-Brown et al., 2007). These conflicting findings are hard to explain if one assumes that masked and overt priming access the same information sources in parallel. Alternatively, it has been proposed that they are sensitive to different stages of word recognition, the former to early pre-lexical access and the latter to a later stage of lemma-level retrieval (Marslen-Wilson, 2007). From this perspective, the apparently conflicting priming results might mean that L1/L2 differences are confined to early stages of form-level access with similar lemma-level processing in L1 and L2. Whilst this possibility remains speculative at present, due to the small number of online studies, examining L2 morphological processing from this perspective could be a promising avenue for future research.

The results from inflection inside derivation also indicate that the two constraints under study affect different stages of processing, both in the L1 and the L2. For L1 English, Cunnings and Clahsen (2008) reported effects of the Category Constraint on early eye-movement measures and of the Structure Constraint on later ones. Similarly, Cunnings and Clahsen (2007) found that constraints on plurals inside compounds also become operative at different points in time during compound processing. With respect to L2 processing of derived words containing inflected base nouns, suppose that the two constraints were applied in parallel. If this was the case, derived words containing singular nouns should yield the shortest reading times (because these forms do not violate any constraint), and derived words containing regular plurals significantly longer ones (because these forms violate both constraints) than those with irregular plurals (due to the violation of just one constraint). Yet, this was not found for the L2 data. Thus, the idea that the two constraints influence L2 processing in parallel was not confirmed by the present data. Instead, the timing of the two constraints was different in L2 processing, with effects of the Category Constraint visible from relatively early eye-movement measures onwards and reliable effects of the Structure Constraint only in the offline data.

2 Is L2 processing simply slower than L1 processing?

Although L1 and L2 processing of morphologically complex words are similar in that specific information sources are consulted at particular moments in time, the question remains as to how the reported L1/L2 differences can be accounted for. One possibility is that the L2 processor makes use of the same mechanisms but that it operates more slowly than the L1 system (e.g. McDonald, 2006).

This idea only provides a partial explanation for the present set of results. It is true that in Experiment 1 response latencies were overall longer in the L2 than the L1 group, as expected from this hypothesis. We also found that, in Experiment 3, effects of the Category Constraint were slightly delayed in the L2 data, recognizable from gaze rather than from first fixation durations onwards, as in the L1 data. Furthermore, reliable effects of the Structure Constraint for the L2 data were only seen in the offline experiment, but (unlike in the L1 data) not for any of the online measures, which may again be indicative of delayed L2 processing. Other findings from the present study cannot be explained in these terms, however. In Experiment 1, even with the modified design which provided extra time for processing the prime word, the L2 group produced the same pattern of results as in the standard design, repetition priming but (unlike the L1 group) no morphological priming, showing that this L1/L2 difference cannot be attributed to slower speed of processing. In Experiment 3, L2 learners responded most sensitively to irregular plural base nouns of derived words on several eye-movement measures, whereas L1 readers had elevated reading times for derived words with regular plural nouns, again a difference that cannot be explained in terms of slowed processing.

3 Are differences in processing due to L1 influence?

Another potential source of L1/L2 differences is the L2 learner's native language(s) which may, for example, cause extra processing effort in cases of conflict with the non-native language. Can the present set of results be explained in these terms?

With respect to the priming results, the experiment reported here replicates results from previous masked priming studies of inflectional morphology with different target languages and across a heterogeneous set of L1 backgrounds. Like the present study, Silva and Clahsen (2008) reported significant priming effects for *-ed* forms in L1 English, but not in L2 learners of English from typologically different L1s. This included German L2 learners of English who have direct equivalents of the English *-ed* in their L1. Yet, in masked priming they showed the same pattern of performance as Chinese and Japanese learners of English, i.e. repetition priming without morphological priming. Neubauer and Clahsen (2009) obtained the same contrast for regular *-t* participles in German, morphological priming in L1 but not in L2 German. Taken together with the present results on Arabic learners of English, these findings suggest that L2 learners' performance in masked priming is not affected by L1 influence.

The potential impact of the L1 on the results of inflection-inside-derivation is harder to assess as we currently do not have any comparable data from L2 learners with L1 backgrounds other than Dutch. In Dutch, *-s* plurals are banned from inside derived words in the same way as in English (**appelsloos* '*applesless'). The results of Experiment 2 which showed the same offline acceptability judgments for the L2 learners as for native speakers of English may therefore perhaps be due to (positive) transfer from the L1. The results from Experiment 3, however, are difficult to explain in these terms. It is unclear how L1 influence could explain why L2 learners' eye movements showed sensitivity to only one of the two relevant constraints, or why L2 learners behaved like English native speakers in the offline but not in the online experiment. Although questions remain with respect to Experiment 2, we conclude that the results on L2 processing are unlikely to be due to L1 transfer.

4 Shallow processing of morphologically complex words

Another possible account for the present dataset comes from the shallow-structure hypothesis (Clahsen and Felser, 2006a). This hypothesis predicts that the L2 comprehension system employs real-time grammatical analysis less than the L1 system and is instead more affected by non-structural (lexical and surface form) properties. We suggest that the results reported here can be explained in this way.

Experiments 1 and 1a tested processes involved in the recognition of morphologically complex words. The results revealed significant masked priming effects for morphologically related prime–target pairs in L1 but not in L2 recognition. The masked priming effect in the L1 has been attributed to repeated stem access, which is made possible by automatic morphological decomposition of the prime word (see, amongst others, Marslen-Wilson, 2007). Thus, an inflected word form such as *prayed* presented as a prime is quickly segmented into its morphological constituents ([[pray]-ed]) thereby making the stem available, hence the reduced recognition time for the target word *pray*, a case of repeated stem access. If one assumes this account for L1 processing, the finding that regularly inflected word forms in a late-learned L2 did not show reliable masked priming effects indicates that these word forms are not automatically decomposed during processing. Thus, in line with the shallow-structure hypothesis, L2 processing relies less on grammatical analysis (in this case, morphological structure) than L1 processing.

Experiment 3 examined the role of lexical and morphological constraints on the processing of derived word with inflected and non-inflected base nouns. The eye-movement data suggest that L2 processing was affected by the constraint against plurals inside derived words in the same way as L1 processing ('Category Constraint'), but not by the morphological constraint that specifically disprefers regular plural base nouns ('Structure Constraint'). To account for these findings, consider the kinds of linguistic information on which the two constraints operate. The Category Constraint prefers a particular stem type, namely bare roots, and excludes inflected words as stems inside derived words. The Structure Constraint, on the other hand, bans stems with a particular internal structure, namely regularly inflected combinatorial forms, from appearing inside derived word forms. Thus, for detecting violations of the Structure Constraint, a deeper level of morphological analysis is required than for violations of the Category Constraint. Whilst the latter can be identified from inspecting the derivational affix (*-less* or *-wise*) and its corresponding stem, detecting violations of the Structure Constraint requires an analysis of the *internal* structure of the base noun within the derived word as to whether it is regularly inflected. The observed differences in sensitivity to these constraints suggest that adult L2 learners do not process complex words at the same level of morphological detail as native speakers and instead rely more on lexical representations, consistent with the shallow-structure hypothesis. In addition, the finding of elevated reading times (specifically first fixation durations) for derived words containing irregular plurals suggests that L2 learners are more sensitive to the surface form of inflected words than to their grammatical structure. Irregular plurals in English have their own word forms which may perhaps make them more easily recognizable as plural forms than those with the unstressed, non-syllabic, non-vocalic, word-final segment *-s*. Thus, irregular plural base nouns may be perceived as striking violations of the Category Constraint, and hence their increased reading times.

V Summary and conclusions

Results from four experiments were presented to investigate morphologically complex words in English. The first two experiments employed the masked priming technique to examine the role of morphological decomposition during early stages of visual word recognition. The first experiment revealed a significant repetition priming effect for a group of advanced Arabic-speaking L2 learners of English in the absence of any reliable morphological priming effect, irrespective of whether target words were presented immediately after prime words or with a temporal delay. Thus, the current study found that providing extra time for processing masked *-ed* forms was not sufficient to obtain morphological priming effects for L2 learners. By contrast, the additional Experiment 1a showed that the masked priming effect for *-ed* forms obtained for L1 English speakers is morphological in nature and not due to either semantic or orthographic relatedness. Experiments 2 and 3 examined derived words containing singular and plural base nouns in high-proficiency Dutch L2 learners of English. Experiment 2 revealed native-like acceptability judgments suggesting that in an offline task, these learners are sensitive to both the Category Constraint against inflected bases for derived words and the morphological Structure Constraint against regular plurals inside derived words. Experiment 3 employed eye-movement monitoring during reading to examine how the same group of L2 learners makes use of these constraints during processing. The results indicated significant effects of the Category Constraint, albeit for later measures than in L1 reading, but no reliable effects of the morphological constraint.

We conclude that L2 processing is not just slower than L1 processing and affected by a learner's native language, but that the L2 comprehension system employs real-time grammatical analysis (in this case, morphological information) less than the L1 system.

Funding

This research was supported by an Alexander-von-Humboldt Professorship to HC and a PhD scholarship of the Kingdom of Saudi-Arabia to LB.

Notes

1. Introducing a delay of 200 ms was motivated by the fact that the mean overall RTs reported by Silva and Clahsen (2008) for the morphologically related condition were on average 224 ms longer for the L2 learners than in the L1 control group.

2. For the error data, the same analyses did not reveal any significant main effects (Condition: $F_1(2, 40) = 1.92$, $p = .16$; $F_2(2,56) = 1.3$, $p = .28$); Prime Type: $F_1(1,20) = 3.37$, $p = .08$; $F_2(1,56) = 1.52$, $p = .22$) or interactions ($F_1(2,40) = 1.5$, $p = .25$; $F_2(2,56) = 1.21$, $p = .30$).

References

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

Anderson S (1992) *A-morphous morphology*. Cambridge: Cambridge University Press.

Baayen H, Piepenbrock R, and van Rijn H (1993) *The CELEX lexical database* [CD ROM]. Philadelphia, PA: Linguistic Data Consortium.

Basnight-Brown D, Chen L, Hua S, Kostić A, and Feldmann L (2007) Monolingual and bilingual recognition of regular and irregular English verbs: Sensitivity to form similarity varies with first language experience. *Journal of Memory and Language* 57: 65–80.

Berent I and Pinker S (2007) The dislike of regular plurals in compounds: Phonological familiarity or morphological constraint? *The Mental Lexicon* 2: 129–81.

Clahsen H and Felser C (2006a) How native-like is non-native language processing? *Trends in Cognitive Sciences* 10: 564–70.

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

Cunnings I and Clahsen H (2007) The time-course of morphological constraints: Evidence from eye-movements during reading. *Cognition* 104: 476–94.

Cunnings I and Clahsen H (2008) The time-course of morphological constraints: A study of plurals inside derived words. *The Mental Lexicon* 3: 149–75.

Di Sciullo A and Williams E (1987) *On the definition of word*. Cambridge, MA: MIT Press.

Fabb N (1988) English suffixation is constrained only by selectional restrictions. *Natural Language and Linguistic Theory* 6: 527–39.

Feldman L, Kostic A, Basnight-Brown D, Filipović-Durdević D, and Pastizzo M (2009) Morphological facilitation for regular and irregular verb formations in native and non-native speakers: Little evidence for two distinct mechanisms. *Bilingualism: Language and Cognition* 13: 119–35.

Forster K and Davis C (1984) Repetition priming and frequency attenuation in lexical access. *Journal of Experimental Psychology: Learning, Memory and Cognition* 10: 680–98.

Gor K (2010) Introduction: Beyond the obvious: Do second language learners process inflectional morphology? *Language Learning* 60: 1–20.

Gordon P (1985) Level-ordering in lexical development. *Cognition* 21: 73–93.

Haskell T, MacDonald M, and Seidenberg M (2003) Language learning and innateness: Some implications of compounds research. *Cognitive Psychology* 47: 119–63.

Jackendoff R (2007) A parallel architecture perspective on language processing. *Brain Research* 1146: 2–22.

Kiparsky P (1982) From cyclic phonology to lexical phonology. In: van der Hulst H and Smith N (eds.) *The structure of phonological representations*. Dordrecht: Foris, 131–75.

Lewis R (2000) Falsifying serial and parallel parsing models: Empirical conundrums and an overlooked paradigm. *Journal of Psycholinguistic Research* 29: 241–48.

Marslen-Wilson W (2007) Morphological processes in language comprehension. In: Gaskel G (ed.) *The Oxford handbook of psycholinguistics*. Oxford: Oxford University Press, 175–93.

McDonald J (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.

Neubauer K and Clahsen H (2009) Decomposition of inflected words in a second language: An experimental study of German participles. *Studies in Second Language Acquisition* 31: 403–35.

Perani D (2005) The neural basis of language talent in bilinguals. *Trends in Cognitive Sciences* 9: 211–13.

Plag I (2002) The role of selectional restrictions, phonotactics and parsing in constraining suffix ordering in English. *Yearbook of Morphology 2001*: 285– 314.

Rastle K, Davis M, Marslen-Wilson W, and Tyler L (2000) Morphological and semantic effects in visual word recognition: A time-course study. *Language and Cognitive Processes* 15: 507–37.

Rastle K, Davis M, and New B (2004) The broth in my brother's brothel: Morpho-orthographic segmentation in visual word recognition. *Psychonomic Bulletin and Review* 1: 1090–98.

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

Seidenberg M, MacDonald M, and Haskell T (2007) Semantics and phonology constraint compound formation. *The Mental Lexicon* 2: 287–312.

Silva R and Clahsen H (2008) Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition* 11: 245–60.

Sinclair J (ed.) (1993) *Collins English dictionary and thesaurus*. New York: Harper Collins.

Stump G (2001) *Inflectional morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.

Ullman M (2005) A cognitive neuroscience perspective on second language acquisition: The declarative/procedural model. In: Sanz C (ed.) *Mind and context in adult second language acquisition: Methods, theory and practice*. Washington, DC: Georgetown University Press, 141–78.

VanPatten B and Jegerski J (eds.) (2010) *Research in second language processing and parsing*. Amsterdam: John Benjamins.

Wiese R (1996) *The phonology of German*. Oxford: Oxford University Press.

Appendix 1 Bio-data of the L2 participants (means; standard deviations in parentheses)

Number of participants	Age (in years)	Age of first exposure to English	Time in the UK (in years)	Proficiency test (maximum score = 100)
Experiment 1: L1-Arabic 20	28 (2.8)	12;8 (0.7)	1;7 (0.9)	79 (4.2)
Experiments 2 and 3: L1-Dutch 25	29;3 (9.5)	9;6 (4.0)	4;8 (5.0)	88.9 (7.3)

Appendix 2 Semantically related and unrelated prime–target pairs for Experiment 1a

Related primes	Unrelated primes	Targets
acquire	foster	GAIN
arrange	arrive	SORT
chase	treat	HUNT
conceal	award	HIDE
dread	trick	FEAR
elevate	pocket	LIFT
finance	garden	FUND
guide	chart	LEAD
hit	vary	BEAT
join	dump	BIND
murder	regard	KILL
nip	dial	BITE
plunge	adjust	DIVE
prepare	ticket	PLAN
quote	breed	CITE
remain	lunch	STAY
rescue	attend	SAVE
seal	jail	PLUG
tempt	urge	LURE
wreck	harm	RUIN
unite	tone	JOIN

Appendix 3 Orthographically related/unrelated prime–target pairs for Experiment 1a

Related primes	Unrelated primes	Targets
carpet	stream	CARP
charge	supply	CHAR
codex	swing	CODE
dollar	minor	DOLL
sight	labour	SIGH
easel	curve	EASE
fellow	custom	FELL
fleece	arcade	FLEE
gateau	butter	GATE
hurtle	strict	HURT
limpid	treaty	LIMP
market	island	MARK
paint	fresh	PAIN
plant	click	PLAN
scarf	castle	SCAR
scant	arrow	SCAN
tackle	brake	TACK
tendon	scream	TEND
twitch	gentle	TWIT
yellow	bottle	YELL
rocket	cement	ROCK

Appendix 4 Frequencies and length of base nouns for the critical items in Experiment 2

	Irregular				Regular		
	Frequency	Letters	Syllables		Frequency	Letters	Syllables
louse	1	5	1	*flea*	4	4	1
ox	4	2	1	*pig*	43	3	1
goose	6	5	1	*duck*	14	4	1
mouse	8	5	1	*rat*	24	3	1
tooth	13	5	1	*eye*	523	3	1
foot	98	4	1	*hand*	724	4	1
woman	338	5	2	*sister*	114	6	2
child	426	5	1	*baby*	258	4	2
man	975	3	1	*brother*	138	7	2
Average	207.67	4.33	1.11		204.67	4.22	1.33

Appendix 5 Frequencies and length of base nouns for the phonological condition in Experiment 2

	Phonologically similar				Phonologically dissimilar		
	Frequency	Letters	Syllables		Frequency	Letters	Syllables
box	102	3	1	*crate*	6	4	1
axe	9	3	1	*spear*	12	3	1
fox	15	3	1	*wolf*	10	4	1
prize	26	5	1	*award*	17	3	2
rose	21	4	1	*daisy*	32	3	2
blouse	11	6	1	*hat*	68	4	1
maize	6	5	1	*corn*	24	6	1
nose	81	4	1	*ear*	88	4	1
Average	33.88	4.13	1.00		32.13	4.25	1.25

Appendix 6 Frequencies and length of base nouns used in Experiment 3

	Irregular plural				Regular plural				Singular		
	Frequency	Letters	Syllables		Frequency	Letters	Syllables		Frequency	Letters	Syllables
lice	2	4	1	*mites*	2	5	1	*flea*	4	4	1
oxen	2	4	2	*animals*	260	7	3	*horse*	132	5	1
geese	5	5	1	*birds*	103	5	1	*fish*	163	4	1
mice	10	4	1	*cats*	67	4	1	*dog*	115	3	1
teeth	75	5	1	*eyes*	523	4	1	*face*	460	4	1
feet	229	4	1	*arms*	210	4	1	*hand*	724	4	1
women	511	5	2	*girls*	438	5	1	*lady*	112	4	2
children	655	8	2	*boys*	349	4	1	*baby*	258	4	2
men	655	3	1	*sons*	206	4	1	*husband*	154	7	2
Average	238.22	4.67	1.33		239.78	4.67	1.22		235.78	4.33	1.33