Neuropsychologia 47 (2009) 158–168

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Word order processing in the bilingual brain

Dorothee Saur[a,*,1], Annette Baumgaertner[b,c,1], Anja Moehring[d,1], Christian Büchel[c], Matthias Bonnesen[d], Michael Rose[c], Mariachristina Musso[a], Jürgen M. Meisel[d]

[a] Department of Neurology, University Medical Center Freiburg, Breisacher Strasse 64, 79106 Freiburg, Germany
[b] NeuroImage Nord, Department of Neurology, University Medical Center Schleswig-Holstein, Schittenhelmstrasse 10, 24105 Kiel, Germany
[c] NeuroImage Nord, Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Martinistr. 52, 20246 Hamburg, Germany
[d] Collaborative Research Center on Multilingualism, University of Hamburg, Max-Brauer-Allee 60, 22765 Hamburg, Germany

ARTICLE INFO

Article history:
Received 16 February 2008
Received in revised form 23 July 2008
Accepted 1 August 2008
Available online 13 August 2008

Keywords:
Syntactic processing
Bilingualism
Age of acquisition
Functional MRI

ABSTRACT

One of the issues debated in the field of bilingualism is the question of a "critical period" for second language acquisition. Recent studies suggest an influence of age of onset of acquisition (AOA) particularly on syntactic processing; however, the processing of word order in a sentence context has not yet been examined specifically. We used functional MRI to examine word order processing in two groups of highly proficient German–French bilinguals who had either acquired French or German after the age of 10, and a third group which had acquired both languages before the age of three. Subjects listened to French and German sentences in which the order of subject and verb was systematically varied. In both groups of late bilinguals, processing of L2 compared to L1 resulted in higher levels of activation mainly of the left inferior frontal cortex while early bilinguals showed no activation difference in any of these areas. A selective increase in activation for late bilinguals only suggests that AOA contributes to modulating overall syntactic processing in L2. In addition, native speakers of French showed significantly higher activation for verb–subject-order than native German speakers. These data suggest that AOA effects may in particular affect those grammatical structures which are marked in the first language.

An important issue in the field of language acquisition research is whether second language acquisition is influenced by a critical period or, more precisely, by maturational changes in the brain (McDonald, 2000; Meisel, 1989, 1997). Although it is commonly accepted that successful first language (L1) acquisition is possible only during a certain age span (Hyltenstam & Abrahamsson, 2003), it is controversial whether the same is true for second language (L2) acquisition. First and second language acquisition exhibit a number of differences with respect to ultimate attainment or learning processes, which have led researchers to postulate a "fundamental difference" between first and second language acquisition (Bley-Vroman, 1989; Meisel, 1991). An influential position claims that there is a critical period in acquiring full competence in two (or more) languages (see Hyltenstam & Abrahamsson, 2003 for a survey of this issue). Because of its emphasis on maturational constraints, this position assumes that age of onset of acquisition (hereafter AOA) is a crucial factor in acquiring full language competence.

There are now a multitude of functional imaging studies using positron emission tomography (PET) and functional magnetic resonance imaging (fMRI) that identify the neural correlates of first and second language processing. Early studies indicated that L1 and L2 share the same basic neural circuits (Chee, Tan, & Thiel, 1999; Hernandez, Martinez, & Kohnert, 2000; Hernandez, Dapretto, Mazziotta, & Bookheimer, 2001; Illes et al., 1999; Perani et al., 1996; Perani et al., 1998). However, L2 processing was found to elicit more activation and more individual variation in activation patterns than L1 processing (Ruschemeyer, Fiebach, Kempe, & Friederici, 2005; Vingerhoets et al., 2003). Increased activation of L2 versus L1 may partly be due to differences in proficiency between the two languages. The less proficient a subject is, the more functional activation has been found; for example in judging semantic relatedness (Chee, Hon, Lee, & Soon, 2001), covert picture naming (De Bleser, 2003), silent word reading (Meschyan & Hernandez, 2006) and past-tense verb identification (Tatsuno & Sakai, 2005). However, Wartenburger et al. (2003) found increased activation of L2 versus L1 even when the levels of proficiency in L1 and L2 were matched. This study is notable in that it varied both proficiency and AOA while examining changes in brain activation in response to grammatically correct sentences and sentences with violations

Abbreviations: L2F, late bilinguals with second language French; L2G, late bilinguals with second language German; 2L1, early bilinguals; La/Lb, the two languages of 2L1 speakers; AOA, age of onset of acquisition.

* Corresponding author. Tel.: +49 761 2705306; fax: +49 761 2705310.
E-mail address: dorothee.saur@uniklinik-freiburg.de (D. Saur).

[1] These authors contributed equally to the paper.

0028-3932/$ – see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2008.08.007

of semantic (selection restriction) and syntactic rules (number, gender or case agreement). Two of the three bilingual groups examined by the authors were highly proficient in both languages (Italian and German); one group had acquired both languages from birth while the other had begun to acquire the L2 after the age of six. When comparing L2 processing between the two groups, late bilinguals showed more activation in inferior frontal gyrus (IFG) bilaterally during syntactic judgements than early bilinguals, while semantic judgements did not result in any differences in activation between early and late highly proficient bilinguals. While semantic processing may be less influenced by AOA, Wartenburger et al. (2003) findings show that grammatical processing is strongly influenced by it.

The aim of Wartenburger et al. (2003) study was to test the claim that syntactic processing in particular is influenced by AOA, compared to other linguistic functions such as semantic processing. The authors compared well-formed canonical sentences with sentences containing violations of gender, number and case in order to assess syntactic processing. In contrast, a recently published study (Hernandez, Hofmann, & Kotz, 2007) used naturally occurring regular and irregular morphology to examine AOA-effects on syntactic processing. The subjects in this study, highly proficient early and late bilinguals, made decisions with respect to the grammatical gender of Spanish nouns. As predicted, Hernandez, Hofmann, and Kotz (2007) found differential activation patterns between the early and late bilingual groups with respect to the processing of irregular (compared to regular) nouns. Specifically, the increase in activity in BA 44/45 for processing irregular (versus regular) items was significantly higher in the late compared to the early bilingual group. The authors suggest that late bilingual subjects used more explicit grammatical rules than early bilinguals for processing irregular items. Thus, additional syntactic processing may be required when processing naturally occurring irregular items in a language that is learned later in life.

The present study was designed to examine whether AOA not only affects only morphology at the single word level (as in Hernandez et al., 2007) but also affects naturally occurring syntactic regularity at the sentence level, such as word order. Previous studies with monolingual subjects have shown that the processing of sentences containing word order violations evokes functional activation in superior and inferior frontal cortex (Embick, Marantz, Miyashita, O'Neil, & Sakai, 2000; Kaan & Swaab, 2002; Newman, 2001). When subjects are asked to judge the grammaticality of sentences made up of pseudo-words, word order violations evoked activation bilaterally in the inferior frontal cortex and the left caudate nucleus and insular cortex (Moro et al., 2001), suggesting that this activation is due to grammatical processing. Importantly, activation predominantly in left IFG has also been observed in *grammatical* sentences with decreasing preference of word order (Friederici, Fiebach, Schlesewsky, Bornkessel, & von Cramon, 2006; Roder, Stock, Neville, Bien, & Rosler, 2002). Thus, changes in functional activation following changes in the order of sentence constituents may be attributed to the linguistic operations underlying the processing of dependency relations between the sentence constituents (Friederici et al., 2006; Roder et al., 2002).

In the present study, sentences had subject–verb (SV) and verb–subject (VS) order, and were either well formed or contained a syntactic violation. Violations were derived from the original (SV and VS) sentences by rearranging sentence constituents while maintaining the original SV or VS sentence structure. Thus, word order in itself could not be used by subjects as a cue to grammaticality because both SV and VS sentence types had an equal likelihood of being either syntactically correct or syntactically incorrect.

In contrast to Hernandez et al. (2007) who included only one L2 group and tested only one language (i.e. Spanish), we included a group of early bilinguals who began to acquire both languages before the age of 3 ('2L1'), and *two* different groups of proficient L2 speakers: a group of subjects who had learned German before learning French (hereafter called 'L2F'), and a group of subjects who had learned French before learning German ('L2G'). In addition to validating an effect of AOA on syntactic processing in L1 and L2 in two independent groups of highly proficient late learners, we were also interested in potential effects of *sequence* of language acquisition on word order processing. Including two groups with complementary L1 and L2 allowed us to directly compare potential constraints on word order processing imposed by the routines of the language learned first. We used French and German because of their differential word order preferences. German exhibits a 'verb-second' word order which means that in main clauses, the verb consistently occupies the second position. French, on the other hand, belongs to the group of 'SVO-languages'. Although French licenses grammatical subjects in a pre- as well as post-verbal position, verb–subject constructions are limited to specific structural contexts, mainly interrogative constructions. Thus, verb-before-subject constructions are the result of a movement process in which the verb moves into a higher functional syntactic category before the subject. With respect to our study, this movement represents the "marked" option, following the definition that markedness, in its general sense, refers to "the presence or absence of a particular linguistic feature" (Crystal, 1997). Hence, the presence of the verb movement (i.e. the VS construction) represents the marked case in French, whereas the absence of this movement (i.e. the SV construction) represents the unmarked case. In addition to the structural differences between German and French, the two languages also differ with respect to the frequency of use of VS constructions. As mentioned above, VS constructions are common in German; the verb precedes the subject whenever some other constituent (an object, WH-word, etc.) is placed sentence-initially (Bayer, 1984; Den Besten, 1977; Platzack, 1983). In colloquial French, post-verbal subjects are infrequent; according to a corpus study VS patterns in sentences with an initial WH-word exhibit a frequency of only 3.6% (Behnstedt, 1973). Based on the above considerations, we hypothesised that an influence of AOA will be more pronounced if processing of the L2 requires syntactic operations that are marked and infrequent in L1. This is because during language acquisition, a marked and infrequent construction in L1 should lead to less developed or less efficient processing routines than a construction which is unmarked and highly frequent.

Based on the finding that grammatical judgments in L2 compared to L1 in late highly proficient bilinguals resulted in more extensive activation predominantly involving Broca's region (e.g. Wartenburger et al., 2003), we hypothesised that the L2-effect would be associated with an increased activation predominantly in IFG. Our prediction with respect to the expected localization of activation in the VS compared to the SV structures was derived from studies with monolinguals that manipulated the word order of the nominal phrases (i.e. subject, indirect object, direct object) in grammatically legal sentences (Friederici et al., 2006; Roder et al., 2002). In both studies, the non-canonical versions, in which the direct (or indirect) object *preceded* the grammatical subject, was rated less acceptable than the canonical (basic) subject–object word order, and resulted in an increase in activation in ventrolateral premotor cortex (area BA 44 (Friederici et al., 2006; Roder et al., 2002)) and prefrontal structures (central and posterior part of the middle frontal gyrus, area BA 6 (Friederici et al., 2006; Roder et al., 2002)). Our study material constitutes a special instance of a non-canonical word order, because in VS-structures, the finite verb is moved out of its (ordinarily) lower position into a higher functional head. Thus, we not only hypothesised that VS order would lead to higher processing costs than SV order in both languages in all three groups, but that VS order would lead to an additional increase in BA 44/BA 6 selectively in the L2G group. This is because VS processing requires syntactic operations that, in addition to being non-canonical, are marked and infrequent in this group's native language, French.

To summarize, the aim of the present study was to examine the neuronal correlate of AOA effects on the processing of word order in early and late bilinguals while keeping proficiency in both languages as comparable as possible. We compared two groups of German/French and French/German bilingual subjects who began to acquire L2 after the age of 10 (L2 speakers) to subjects who began to acquire both languages before the age of three (2L1 speakers). Early bilingual rather than monolingual speakers were included in order to be able to validly compare AOA effects while controlling for potential mutual influences of the respective languages. To the extent possible, language proficiency was matched between groups in order to minimize potential confounding influences of differences in proficiency on functional activation in early and late bilinguals.

We expected to find (i) overall increased functional activation in the two L2 groups for grammatical processing in L2 versus L1, predominantly in areas previously associated with syntactic processing in highly proficient late bilinguals (i.e. left frontal cortex comprising posterior IFG), while we did not expect differences between processing La and Lb in the 2L1 group, and (ii) selectively increased functional activation for the processing of VS- compared to SV-word order in the L2G compared to the L2F group, particularly in those areas that have previously been associated with higher demands on word order processing in monolinguals (i.e. BA 44 extending into BA 6). This expectation is based on the fact that for the L2G group, VS sentence structures are marked.

1. Materials and methods

1.1. Subjects

Thirty-six healthy, right-handed bilingual (German–French) adults participated. They were assigned to three groups according to their AOA: an L2F-group ($n = 12$, three males, second language French, $AOA_{L2} > 10$ years), an L2G-group ($n = 12$, one male, second language German, $AOA_{L2} > 10$ years) and a 2L1-group ($n = 12$, two males, both languages (La, Lb) were acquired <3 years). All subjects were living in Germany when they participated in the study. All participants were either graduating students or graduates, i.e. they had a comparable educational background. Right-handedness was tested with the 10-item version of the Edinburgh Handedness Inventory (Oldfield, 1971). Proficiency was measured with two multiple choice proficiency tests. The administration of each test took approximately 30 min. For German we used a composite of two accredited proficiency tests (Eichheim & Storch, 1992; Goethe-Institut, 1981) with 61 items; for French we used the REEKS Test with 88 items (developed and used at the Institute for Modern Languages of the Katholieke Universiteit Leuven, not published). Only those subjects were included who demonstrated high proficiency in both languages, as defined by a proficiency score of more than 70% in both languages. Demographics and language background of the subjects are given in Table 2A. All subjects gave written informed consent prior to the investigation and were paid for participation. The study was approved by the local Ethics Committee of the University Medical Center Hamburg-Eppendorf.

1.2. Stimulus material

Forty original sentences in each language were used to derive the syntactically correct and incorrect versions of the SV and the VS conditions, respectively, resulting in four stimulus categories with 40 items each (see Table 1 for stimulus examples). In about half of the cases, the construction of correct and incorrect versions of an SV (or VS) sentence type required a modification of the original sentence. Consequently, stimuli consisted of both statement and question sentences. Prosody did not provide sufficient cues for rating grammaticality, since WH-questions occurred in both, correct and incorrect, stimulus types.

In total, there were 160 German and 160 French sentences. Three native speakers of French and German each judged stimuli with respect to their categorization as grammatically correct or incorrect versions of the SV and VS sentence types. Sentences were constructed by using highly frequent words in German and French (Juilland, Brodin, & Davidovitch, 1970; Ruoff, 1990). Sentences were spoken by a female bilingual speaker. The intonation of all stimuli was carefully controlled to be as natural as possible.

1.3. Experimental design and task

The study used a mixed design, with group (L2F, L2G, 2L1) as a between subjects variable, and word order (subject–verb or verb–subject) and sentence acceptability (grammatically correct versus incorrect) as within-subject variables. Stimulus sentences were divided into eight sessions with 40 sentences each (10 sentences of each category), with the restriction that the correct and incorrect version of each stimulus sentence could not occur in the same session. Within a session, either German (G) or French (F) sentences were presented (order of the sessions: GFGFFGFG or FGFGGFGF). The sequence of sentences within sessions was pseudo-randomised with respect to subject–verb versus verb–subject order and correct versus incorrect word order. Specifically, no more than two sentences of the same type, and no more than three correct or incorrect sentences could occur in a row. To increase design efficiency, each session contained eight silent (null) events which were randomly interspersed among the 40 sentences. Duration of sentences ranged from 1360 to 3100 ms; the interstimulus interval varied between 2900 and 4600 ms resulting in a trial duration of about 6000 ms.

Before the scanning session, the experiment was explained and the subjects had a brief training session outside the scanner. Conversation before and during scanning was in German. The sentences were presented aurally using MR-compatible headphones. Subjects were asked to indicate by left-handed button-press whether they had identified a correct or incorrect sentence. The start of each session was announced by brief instructions via headphones. During scanning, subjects kept their eyes open.

1.4. MR scanning

Functional MRI was performed on a 3T Siemens TRIO system (Siemens, Erlangen, Germany). A total of 135 fMRI volumes per session with 36 contiguous axial slices covering the whole brain (3 mm thickness, 1 mm gap) was acquired using a gradient echo echo-planar (EPI) T2*-sensitive sequence (TR 2.24 s, TE 30 ms, flip angle 80°, matrix 64×64 pixel^2). The first five volumes were discarded to allow for T1 equilibration effects.

1.5. Data analysis

On the behavioural data we performed a one-factorial ANOVA and post hoc two-sample *t*-tests (two-sided) to test for significant differences in proficiency/task performance between groups; within groups we tested for significant differences in French-/German sentence-processing using paired *t*-tests (two-sided).

Imaging data were analysed using Statistical Parametric Mapping [SPM2; Wellcome Department of Imaging Neuroscience (Worsley & Friston, 1995)] implemented in Matlab 6.5 (Mathworks, Natick, MA, USA).

Pre-processing: All slices were corrected for different acquisition times of signals by shifting the signal measured in each slice relative to the acquisition of the

Table 1
Sentence material with four categories for each language (with literal translation)

	German	French
Subject–verb-order (SV)		
Correct	**Peter (S) kommt (V) spät von der Arbeit**	**Natalie (S) travaille (V) á Paris ce soir**
	(*Peter comes home late from work*)	(*Natalie works in Paris tonight*)
Incorrect	*Spät Peter (S) kommt (V) von der Arbeit	*Où Natalie (S) travaille (V) ce soir?
	*(*Late Peter comes home from work*)	*(*Where Natalie works tonight?*)
Verb–subject-order (VS)		
Correct	Wann kommt (V) Peter (S) von der Arbeit?	Où travaille (V) Natalie (S) ce soir?
	(*When comes Peter home from work?*)	(*Where works Natalie tonight?*)
Incorrect	*Spät kommt (V) von der Arbeit Peter (S)	*Ce soir travaille (V) Natalie (S) á Paris
	*(*Late comes home from work Peter*)	*(*Tonight works Natalie at Paris*)

Note: original sentences are written in bold letters, * indicates violation of word order.

middle slice. All volumes were then spatially realigned to the first volume in order to correct for movement. Resulting volumes were spatially normalised to a standard image template based on the Montreal Neurological Institute (MNI) reference brain, and resampled to 3 mm × 3 mm × 3 mm voxel size (Friston, Ashburner, Frith, Poline, & Frackowiak, 1995b). All normalised images were then smoothed using an isotropic 12-mm Gaussian kernel to account for inter-subject differences. Statistical analysis: At the first level, the four conditions (subject–verb-order correct/incorrect, verb–subject-order correct/incorrect) for each language (German/French) were modelled as eight separate conditions. Sentence onsets and durations were convolved with a canonical hemodynamic response function as implemented in SPM2. Only correctly judged items were included in the analysis. Voxel-wise regression coefficients for all conditions were estimated using least squares within SPM2 (Friston et al., 1995a).

Our research questions were addressed in a second-level analysis treating volunteers as a random effect, in which the contrast images of the eight conditions were entered into an ANOVA. Correction for non-sphericity resulting from unequal variances between the three groups was implemented. With respect to our first research question, differences between L1 and L2 grammatical processing were investigated in the three groups. Separately for all three groups, we computed the main effects of both languages and the differential effects between the languages by pooling over all German and French conditions, irrespective of whether or not sentences were grammatical. To identify common areas of L2-grammatical processing in the L2G- and L2F-group, we performed a conjunction analysis under the conservative conjunction null hypothesis (Nichols, Brett, Andersson, Wager, & Poline, 2005) using the L2 > L1 contrasts of the L2F- and L2G-groups computed for the whole brain. To quantify the observed effects, we extracted the individual parameter estimates of German and French language processing at the peak voxel of each activated cluster identified in the conjunction analysis. Parameter estimates of the 2L1-group were included for comparison. Statistical comparisons of the extracted data within the peak voxels were done using SPSS 13.0 software. To test for significant differences across conditions, we performed an ANOVA on the six-parameter estimates, separately for each area. When a significant result was obtained, post hoc paired *t*-tests (two-sided) were carried out within groups to test for significant differences between processing in L1 and L2.

With respect to our second research question, we analysed the word order effect (subject–verb versus verb–subject) separately in each group. To avoid potential confounds by error detection, re-analysis or repair processes, we analysed only the grammatically correct sentences (which were correctly judged). Again, parameter estimates for the SV- and VS-conditions of each group were computed in the peak voxel of each activated area by extracting the data from the respective contrast images of each subject. Statistical analysis was done as described above.

Levels of significance were set as follows: behavioural data and parameter estimates $p < 0.05$; fMRI data $p < 0.05$ corrected for multiple comparisons (exceptions are indicated in the legends). Since the extent (i.e. the cluster size) of activation is strongly dependent on the statistical threshold used, we use the term 'higher activation' only if the statistical contrasts yield significant increases in the strength of activation, at the voxel level.

2. Results

2.1. Behavioural results

The analysis of the demographic data showed a higher average age in the L2G than in the other two groups ($F(2, 33) = 9.2$, $p < 0.05$). The difference in AOA of L2 in the L2F- and L2G-group was not significant ($T = 1.0$; $p = 0.32$), whereas time of total exposure in the "L2-country" was longer for L2G compared to L2F ($T = 3.4$, $p < 0.05$).

2.1.1. Proficiency

(i) *Differences between groups*: Proficiency differed for German ($F(2, 33) = 25.6$, $p < 0.05$) and French ($F(2, 33) = 28.3$, $p < 0.05$) between the three groups. For German, the L2G group showed a significantly lower proficiency compared to the L2F- and 2L1-group. For French, the L2F-group showed a significantly lower proficiency compared to the L2G-group. However, the 2L1-group also showed significantly lower proficiencies in French and German compared to the respective L1's of the L2G- and L2F-groups (Table 2A, Fig. 1). (ii) *Differences within groups*: The proficiency of L2 compared to L1 was significantly lower in the L2F- and L2G-group; however, the 2L1 group also showed a significantly lower proficiency in French compared to German.

2.1.2. Task performance during scanning (German/French)

(i) *Differences between groups*: The three groups showed no significant differences in task performance neither for French (F(2,

Table 2A
Subjects and behavioural data

	L2F	L2G	2L1	Differences between groups (one-factorial ANOVA; post hoc 2 sample *t*-tests)
n	12	12	12	
Sex				
f	11	9	10	
m	1	3	2	
Age (M, S.D.) (years)	29.3 (6.3)	41.6 (14.0)	26.2 (5.0)	
Age of acquisition (M, S.D.) (years)				
G	Since birth	15.6 (5.7)	Before 3	
F	13.9 (2.4)	Since birth	Before 3	
Time of total exposure to L2 (M, S.D.) (years)				
G	–	13.2 (11.6)	–	
F	1.8 (1.2)	–	–	
Proficiency testing (mean of correct items in % (S.D.))				
G	98.4 (0.7)	86.4 (5.7)	94.9 (3.3)	**$F(2, 33) = 25.6$, $p < 0.001$** **L2F vs L2G ($T = 6.9$, $p < 0.05$)** **L2F vs 2L1 ($T = 3.4$, $p = 0.05$)** **L2G vs 2L1 ($T = 4.3$, $p < 0.05$)**
F	84.7 (5.9)	98.3 (0.6)	85.3 (6.9)	**$F(2, 33) = 28.3$, $p < 0.001$** **L2F vs L2G ($T = 7.9$, $p < 0.05$)** L2F vs 2L1 (T = 0.2, p = 0.8) **L2G vs 2L1 ($T = 6.5$, $p < 0.05$)**
Differences within groups (paired *t*-test)	**$t = 7.5$, $p = 0.05$**	**$t = 6.7$, $p = 0.05$**	**$t = 4.3$, $p < 0.05$**	
Task performance during scanning (mean of correct items in % (S.D.))				
G	91.0 (4.6)	85.9 (6.9)	86.7 (7.1)	$F(2, 33) = 2.5$, $p = 0.1$
F	74.5 (8.4)	81.3 (6.4)	78.2 (3.6)	$F(2, 33) = 2.6$, $p = 0.09$
Differences within groups (paired *t*-test)	**$t = 6.5$, $p < 0.05$**	$t = 1.9$, $p = 0.74$	**$t = 2.6$, $p = 0.03$**	

Note. f, female; m, male; M, mean; S.D., standard deviation; F, French; G, German; L2F, late bilinguals with second language French; L2G, late bilinguals with second language German; 2L1, early bilinguals. Boldface indicates significance at $p < 0.05$.

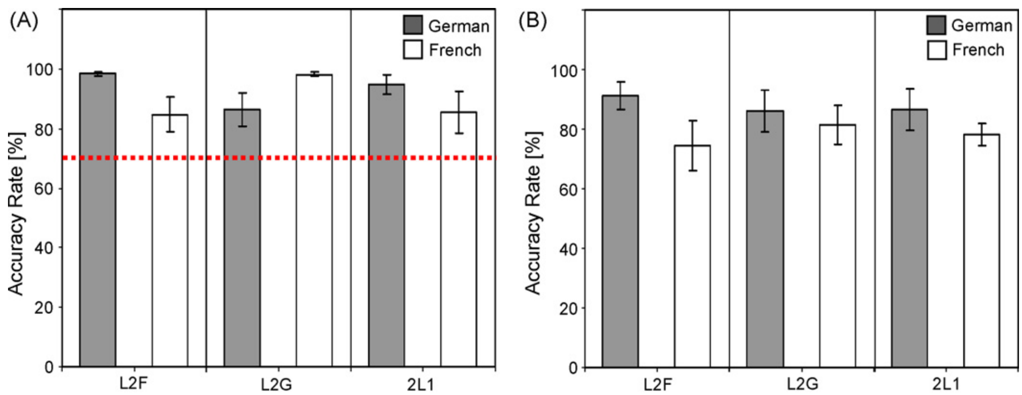

Fig. 1. Behavioural data. Plots of the accuracy rate in the proficiency tests (A) and in the task during scanning (B) for German and French stimuli in the three groups (both in % correct ratings). Mean and 2× S.D. are displayed, red line indicates cut off for inclusion into the study (70% correct items in the proficiency testing).

33) = 2.5, $p = 0.1$) nor for German ($F(2, 33) = 2.6$, $p < 0.09$). (ii) Differences within groups: The L2F group showed a significantly lower task performance for French compared to German, whereas in the L2G and 2L1 group no significant differences in error rate between both languages were observed (Table 2A, Fig. 1).

2.1.3. Task performance during scanning (SV-/VS-condition)

The analyses of the SV- and VS-condition included only the grammatically correct items. (i) *Differences between groups*: there were no significant differences between the three groups in any of the four conditions [SV-/VS-condition in both German and French ($p > 0.32$)]. (ii) *Differences within groups*: In the L2F-group, task performance differed between the four conditions ($p = 0.009$). Post hoc paired *t*-tests in this group showed a significantly better task performance in the SV- compared to the VS-condition in German ($p = 0.004$) and French ($p = 0.02$). In the L2G and 2L1 groups, task performance on French items was better in the SV-condition; however, the difference between the SV- and the VS-condition was not significant in any of the two groups (L2G group: $p = 0.23$; 2L1 group: $p = 0.26$; Table 2B)).

2.2. Functional neuroimaging

2.2.1. Main effects of German and French

Comparison of German and French stimuli with the resting state revealed broad bilateral activation mainly within the temporal and frontal lobe with highest activation in primary auditory areas of the left and right superior temporal gyrus. This pattern was similar in all three groups for German and French (no Figure or Table).

2.2.2. Comparison of L2 and L1

Processing of L2 compared to L1 resulted in increased activation in several brain regions. In the L2F-group, processing L2 (French) compared to L1 resulted in significantly higher activation in the left IFG, bilaterally in the caudate nucleus, and in the left middle/inferior temporal gyrus. The L2G group showed significantly higher activation in L2 (German) compared to L1 in the left IFG, anterior cingulate, the right caudate nucleus and the left superior occipital cortex. However, there were also differences between both languages in the 2L1 group. Higher activation in German (La) compared with French (Lb) was found bilaterally in the insular cortex, the right Cerebellum and SMA; higher activation in French compared with German was found in the left fusiform gyrus and in the IFG (results of all comparisons are listed in Table 3 and displayed in Fig. 2).

2.2.3. L2-effect

To compute conjoint effects of L2-language processing in the two late bilingual groups, we performed a conjunction analysis with the L2 > L1 comparisons of both groups (i.e. $L2F_{L2>L1}$ and $L2G_{L2>L1}$). This analysis showed higher activation for L2-sentence processing in the pars opercularis and pars triangularis of the left IFG which extended into BA 6, in the right caudate nucleus and in the left inferior temporal gyrus (Table 3 and Fig. 2). In the peak voxels of these clusters, we computed the parameter estimates for German and French sentence processing of all three groups. The ANOVA of the extracted estimates showed that the size of the effect was significantly different across conditions for German and French sentence processing in all four areas (all four $F(2, 141) > 7.7$; $p < 0.001$). As expected from the conjunction analysis, the identified voxels were activated by both languages. However, in the two late bilingual groups, activation in the respective L2 was significantly higher than in the respective L1 in each of the four areas (post hoc paired *t*-tests, $T > 3.1$, $p < 0.05$). In contrast, the 2L1 group showed no significant differences between the two languages La and Lb in all identified areas (all four post hoc paired *t*-test, $T < 0.8$, $p > 0.4$; Fig. 3A–D).

2.2.4. L1-effect

We also computed the reverse conjunction of L1 > L2 in the L2F- and L2G-group (i.e. $L2F_{L1>L2}$ and $L2G_{L1>L2}$). This analysis showed no significant result, even after lowering the statistical threshold to $p < 0.01$, uncorrected for multiple comparisons.

Table 2B
Task performance during scanning for the four conditions (only grammatically correct items)

	Subject–verb-order		Verb–subject-order		Differences within group
	German	French	German	French	
L2F (M, S.D.)	93.5% (3.06)	88.1% (3.14)	89.0% (3.53)	81.0% (4.72)	**$F(3, 40) = 4.4$, $p = 0.009$**
L2G (M, S.D.)	89.8% (2.59)	89.5% (5.86)	89.5% (3.22)	83.4% (3.04)	$F(3, 40) = 1.2$, $p = 0.32$
2L1 (M, S.D.)	88.1% (4.27)	87.7% (1.88)	87.9% (3.56)	84.8% (3.5)	$F(3, 40) = 0.4$, $p = 0.79$
Differences between groups	$F(2, 31) = 1.3$, $p = 0.32$	$F(2, 31) = 0.1$, $p = 0.91$	$F(2, 31) = 0.09$, $p = 0.91$	$F(2, 31) = 0.4$, $p = 0.65$	

Task performance during scanning for the four conditions in % of grammatically correct items, M, mean; S.D., standard deviation; Boldface indicates significance at $p < 0.05$.

Table 3
Comparison of German vs. French

Region	Side	MNI			t
		x	y	z	
(1) L2F (German > French)					
Precentral gyrus	R	39	−21	57	8.44
Supramarginal gyrus	R	57	−21	21	8.23
Insular cortex	R	45	3	9	7.83
Cerebellum	L	−27	−45	−30	7.78
Insular cortex	L	−42	0	9	7.13
SMA	M	0	−3	51	6.57
Supramarginal gyrus	L	−57	−24	24	5.46
(2) L2F (French > German)					
Inferior frontal gyrus, pars triangularis	L	−45	15	21	7.23
Caudate nucleus/putamen	R	15	12	−3	6.78
Caudate nucleus	L	−12	9	0	6.19
Middle temporal gyrus	L	−63	−54	0	6.12
(3) L2G (German > French)					
Inferior frontal gyrus, pars opercularis	L	−54	12	3	7.17
Cingulum	R	12	18	33	6.15
Caudate nucleus	R	12	21	9	5.73
Superior occipital cortex	L	−24	−78	39	5.06
(4) L2G (French > German)					
Fusiform gyrus	R	42	−12	−33	5.72
(5) 2L1 (German > French)					
Insular cortex	L	−36	−6	12	11.35
Insular cortex	R	42	3	6	10.66
SMA	R	6	−3	51	7.81
Cerebellum	R	−24	−54	−24	7.23
(6) 2L1 (French > German)					
Fusiform gyrus	L	−48	−51	−12	5.24
Inferior frontal gyrus, pars triangularis	L	−51	27	15	4.58
Conjunction: $L2G_{German>French}$ and $L2F_{French>German}$					
(A) Inferior frontal gyrus, pars opercularis	L	−48	12	18	5.03
(B) Inferior frontal gyrus, pars triangularis	L	−45	33	6	4.07
(C) Caudate nucleus	R	12	15	0	4.72
(D) Inferior temporal gyrus	L	−48	−51	−15	4.24

Peak voxels for the within group contrasts L2 > L1 and L1 > L2 at a statistical threshold of $p < 0.05$ ($t > 4.53$), corrected for multiple comparisons across the whole brain. For the conjunction analysis the statistical threshold was set a $p < 0.001$ ($t > 3.13$). MNI refers to Montreal Neurological Institute reference brain coordinates; R right, L left.

2.2.5. Word order effect (grammatically correct sentences only)

To examine whether word order processing is modulated by the sequence of acquisition, we directly compared the grammatically correct SV- and VS-sentences in all three groups. In these analyses, German and French stimuli were considered together. In all groups, there was higher activation in the VS-condition mainly in the left inferior and middle frontal cortex, the inferior parietal lobe and the cerebellum, while the opposite comparison did not activate these

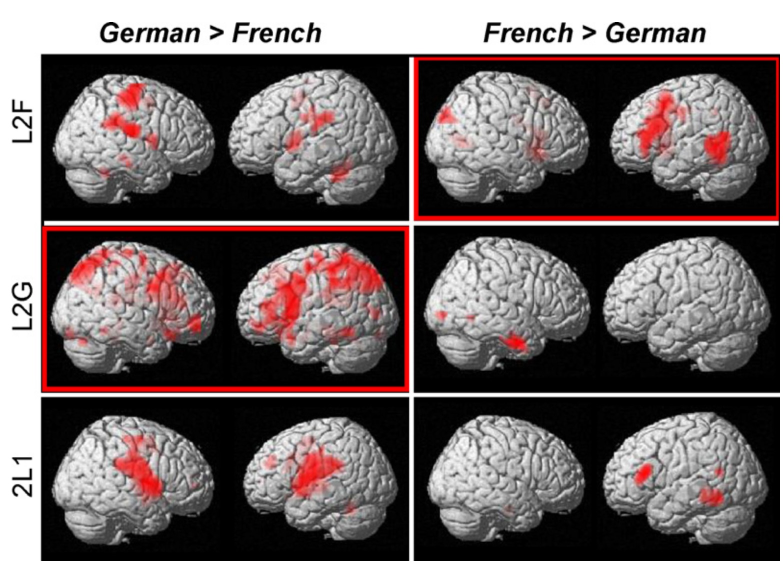

Fig. 2. fMRI activation patterns for German and French. Differential effects of German vs. French and vice versa ($t > 4.53$, $p > 0.05$ corrected for multiple comparisons) are rendered on a canonical brain. Contrasts with a red frame constitute the input into the conjunction analysis shown in Fig. 3.

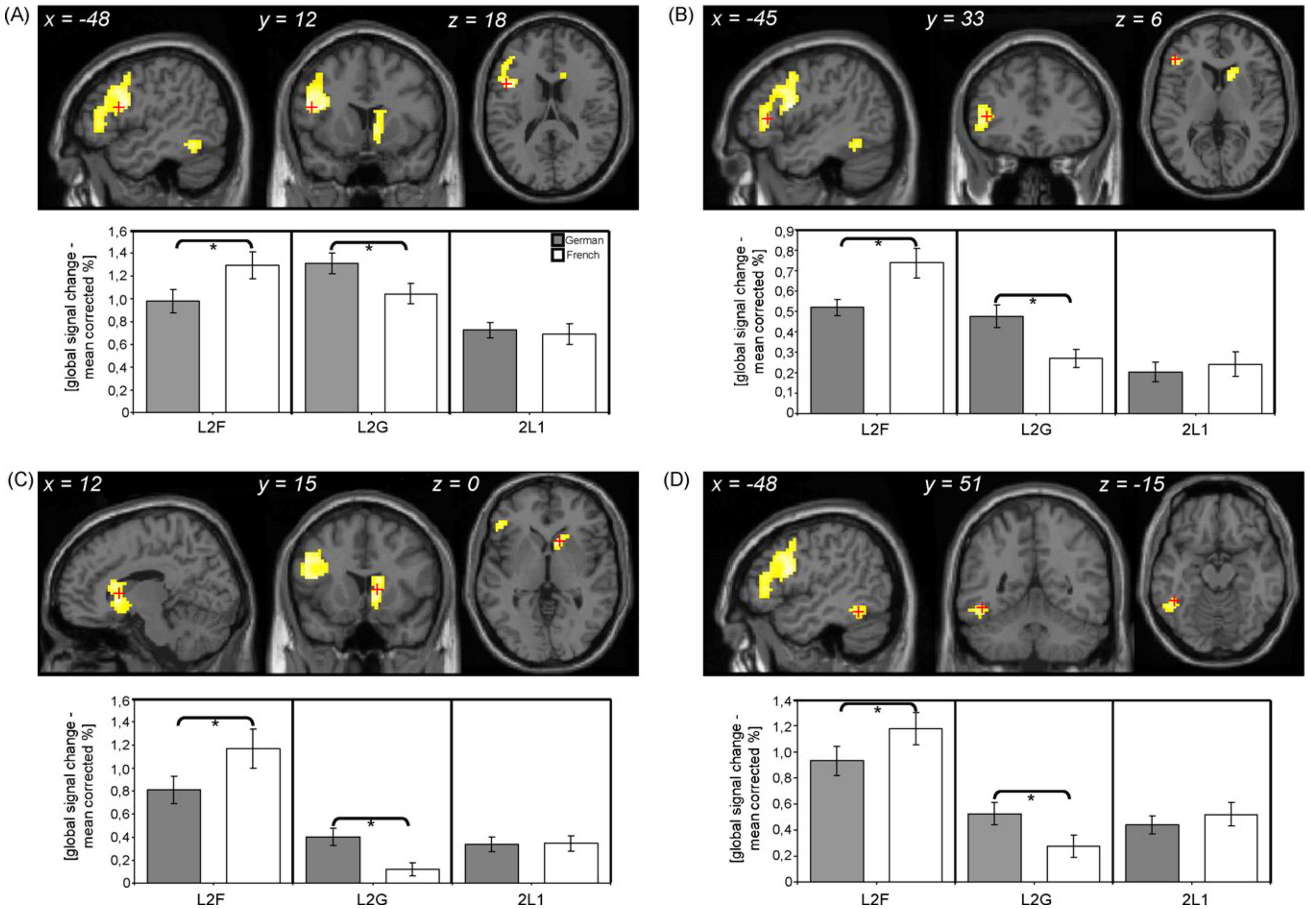

Fig. 3. Results of the conjunction analysis of $L2F_{L2>L1}$ and $L2G_{L2>L1}$. All voxels are significant at $p<0.001$ (uncorrected for multiple comparisons, $t>3.13$). Peak voxels are indicated with a red + and located in the left IFG, pars opercularis (A) and pars triangularis (B), the right head of the caudate nucleus (C) and the left inferior temporal gyrus (D). Plots represent mean corrected global signal change in the indicated voxels for French (F) and German (G) sentence processing in late French speakers (L2F), late German speakers (L2G) and early bilingual subjects (2L1). *Indicates significant differences ($p<0.05$) within groups between German and French sentence processing in a paired t-test (two-sided); error bars represent the 95% confidence interval.

areas (Fig. 4A, Table 4). According to the premise that VS-order is a marked structure in French, we examined whether the processing of VS-order leads to higher activation in the L2G-group compared with the L2F-group (interaction $L2G_{VS>SV}>L2F_{VS>SV}$). The analysis revealed a single cluster in the left inferior frontal cortex at the border of BA 44 to BA 6. The parameter estimates for the peak voxel in this cluster show significantly higher activation in the VS-condition exclusively for the L2G group, while this comparison is

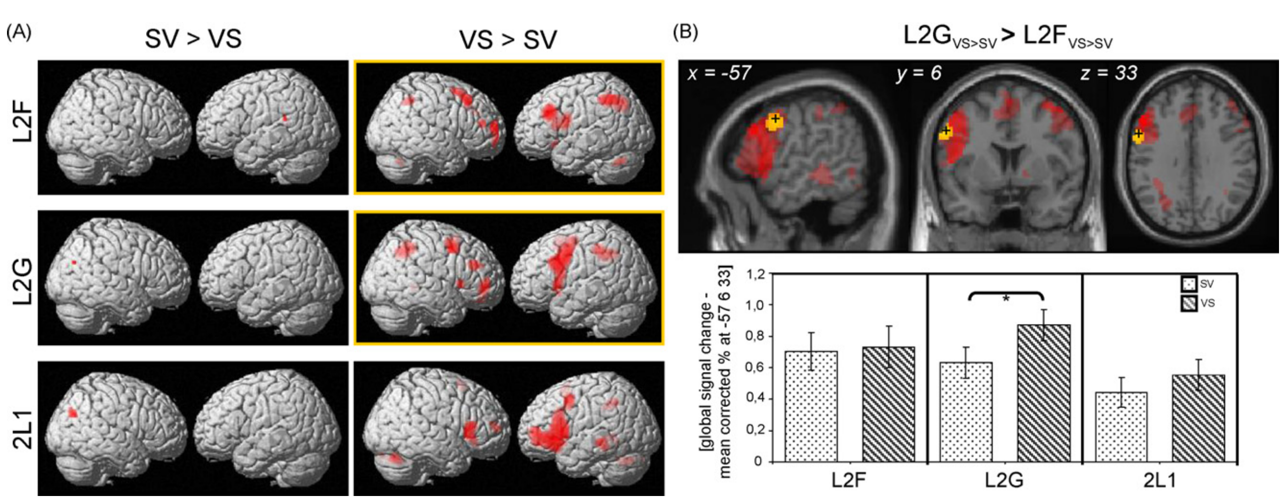

Fig. 4. Word order effect (subject–verb (SV) and verb–subject (VS)) in the three groups. (A) Word order effect SV > VS and VS > SV in the three groups, $p<0.001$, uncorrected for multiple comparisons. Contrasts with a yellow frame constitute the input into the interaction analysis shown in Fig. 4B (yellow). (B) Word order effect in the interaction analysis $L2G_{VS>SV}>L2F_{VS>SV}$ is indicated in yellow ($p<0.01$, uncorrected for multiple comparisons). For direct comparison, the word order effect (VS > SV) of all three groups is displayed in red ($p<0.001$, uncorrected for multiple comparisons). Analyses in A and B contain only the grammatically correct German and French stimuli.

Table 4
Word order effect for grammatically correct German and French stimuli

Region	Side	MNI			t
		x	y	z	
(1) L2F: VS > SV					
Middle frontal gyrus	L	−48	27	33	4.39
Inferior parietal lobe	L	−39	−48	48	3.89
Middle frontal gyrus	R	45	24	48	3.57
Cerebellum	L	−30	−63	−33	3.48
(2) L2F: SV > VS					
–	–	–	–	–	–
(3) L2G: VS > SV					
Inferior frontal gyrus, pars opercularis	L	−48	6	33	4.40
Inferior parietal lobe	L	−36	−42	42	4.08
Inferior parietal lobe	R	36	−54	42	3.72
Inferior frontal gyrus, pars opercularis	R	54	18	3	3.08
SMA	R	6	12	57	3.04
(4) L2G: SV > VS					
–	–	–	–	–	–
(5) 2L1: VS > SV					
Inferior frontal gyrus, pars triangularis	L	−51	42	−3	4,76
Cerebellum	R	36	−69	−30	3.63
Middle temporal gyrus	L	−48	−33	−6	3.60
Inferior frontal gyrus, pars triangularis	R	60	21	3	3.49
Inferior parietal lobe	L	−30	−48	45	3.29
Cerebellum	L	−27	−66	−33	3.09
(6) 2L1: SV > VS					
Angular gyrus (BA 39)	R	51	−72	30	3.36
(7) Interaction: $L2G_{VS>SV}$ > $L2F_{VS>SV}$					
Precentral gyrus (BA 6) bordering BA 44	L	−57	6	33	2.49

Statistical threshold was set at $p < 0.001$ ($t > 3.13$).

not significant for the L2F and 2L1 group (Fig. 4B). The reverse interaction of $L2F_{VS>SV}$ > $L2G_{VS>SV}$ did not show significant results at any statistical threshold.

3. Discussion

This study investigated grammatical processing with fMRI in three groups of highly proficient bilingual subjects: a first group of native German speakers who acquired French after the age of 10, a second group of native French speakers who had acquired German after the age of 10, and a third group of early bilinguals who had learned and practiced both languages before the age of three. We presented German and French syntactically correct sentences and sentences with a syntactic violation created by shifting or substituting sentence constituents while maintaining the original subject–verb or verb–subject word order.

Although the three groups did not differ in task performance, L2 speakers' activation patterns differed systematically from those of early bilinguals. Both groups of late bilinguals independently showed a significantly higher level of activation during grammatical processing in L2 than in L1 in the left inferior and middle frontal gyrus including the pars triangularis (BA 45) and opercularis (BA 44), in left inferior temporal gyrus and in the basal ganglia.

The highest level of BOLD signal change in both L2 groups was observed in the left IFG (pars opercularis (BA 44) and pars triangularis (BA 45), and extending into dorsolateral prefrontal cortex). Left IFG activation, predominantly within BA 45 and BA 44, has been associated with monolingual syntactic processing in numerous functional imaging studies using syntax production (e.g. Haller, Radue, Erb, Grodd, & Kircher, 2005) as well as syntactic comprehension tasks on well-formed sentences (e.g. Caplan, Alpert, & Waters, 1999; Dapretto & Bookheimer, 1999; Embick et al., 2000). Activation in left IFG typically increases with grammatical complexity (e.g. Caplan, Alpert, Waters, & Olivieri, 2000; Just, Carpenter, Keller, Eddy, & Thulborn, 1996; Wartenburger et al., 2004). Also, left IFG (Broca's area) gets more involved as word order in a grammatical sentence becomes less acceptable (Roder et al., 2002). IFG activation has also been observed when ungrammatical sentences are compared with grammatical sentences ((Ni et al., 2000; Wartenburger et al., 2004); but see (Friederici et al., 2006), who found specific activation increases for syntactic violations in the deep frontal operculum)). Similar findings exist for bilinguals. In proficient bilinguals, more extensive activation was found in IFG when judging number, gender or case agreement in L2 than in L1 (Wartenburger et al., 2003). Particularly pertinent to our findings is the study by (Golestani et al., 2006) in which subjects either silently read sequences of three to five words, or covertly generated simple but grammatically complete sentences from the same visually presented three to five words. Sentence generation compared to word reading resulted in pronounced left IFG activation in BA 44/45 which was stronger for L2 (English) than for L1 (French). While Golestani et al. (2006) caution that the contrast of sentence generation with word reading may encompass semantic, lexical, and phonological processes beyond the intended (syntactic) sentence building processes, our finding of predominantly BA 44/45 activation for word order processing in L2 compared to L1 in two independent groups of proficient late bilinguals strikingly coincides with and corroborates Golestani et al. (2006) results.

Our results show a specific involvement of left posterior inferior frontal cortex in word order processing in L2 compared to L1. Since left IFG has often been discussed as a core area mediating verbal working memory, the question arises whether the observed increase in activation for L2 compared to L1 processing may be explained by higher working memory demands imposed by the processing of word order in L2. This question is particularly pertinent considering that word order processing involves sequentially structured events, which have been postulated to be preferentially processed in left BA 44/45 (Schubotz & von Cramon, 2004). In using a novel paradigm in which subjects performed syntactic judgement and memorization tasks with the same set of sentences, Hashimoto and Sakai (2002) were able to dissociate verbal working memory from syntactic processing proper, which was centered in IFG. A study by Caplan et al. (2000) used a different route to examine whether subvocal rehearsal may explain the increase in activation in left IFG when processing grammatically more complex sentences. An increase in IFG with increasing syntactic complexity was observed even when subjects articulated the word "double" while making grammatical judgments, in order to inhibit their ability to internally rehearse the sentences. These findings suggest that even though working memory functions may partly be involved in syntactic processes, they are not able to entirely explain the left IFG activation in response to syntactic processing tasks.

Apart from the large cluster of activation in left inferior frontal cortex, we found that in the two L2 groups, word order processing in L2 (versus L1) was also associated with increased activation in the basal ganglia. Increased activation in the basal ganglia during grammatical processing in L2 versus L1 is consistent with previous findings in highly proficient late bilinguals (e.g. Wartenburger et al., 2003). Other recent functional imaging studies which similarly presented stimuli in two different languages also showed an involvement of basal ganglia structures. In a recent study at the word level, proficient bilinguals made semantic decisions with respect to the visual features of objects denoted by written words, which were preceded by written words which served as primes and which subjects were asked to ignore (Crinion et al., 2006). When the prime and target words were semantically related and when they were presented in the same language, there was reduced activation in the left caudate nucleus. This lead the authors to propose that the (left) caudate nucleus plays an important role in language control. This conclusion is supported by another recent finding using auditory narratives. When subjects were passively listening to narratives, there was increased activation in the basal ganglia when the language of the narrative (unpredictably) switched between L1 and L2 (Abutalebi et al., 2007). Together, these findings suggest that basal ganglia activation may be related to cognitive control mechanisms and may reflect the cost of switching between languages (but see van Heuven, Schriefers, Dijkstra, & Hagoort, 2008, who did not find basal ganglia activation in a lexical decision task with stimuli that induced a language conflict). Since French and German sentences occurred equally often in our study, and since subjects had to frequently switch between French and German processing modes (especially because instructions were given in German), an increase in activation for L2 compared to L1 stimuli may reflect processes of language control and language selection rather than linguistic processes per se.

Activation in left inferior frontal cortex for grammatical processing in L2 (versus L1) extended superiorly into BA 6 in both groups of late bilinguals. Increased activation in middle frontal areas may reflect increased cognitive effort in performing grammaticality judgements in L2 (Badre & Wagner, 2004; Gabrieli, Poldrack, & Desmond, 1998). However, this does not seem to be an exhaustive explanation since late speakers of German performed equally well (or better) on German than on French items on the judgement task. A complementary explanation may be provided by the assumption that grammatical processing is affected by maturational changes in the brain. Conceivably, as the brain settles on specific (grammatical) parameters during the critical period for language acquisition, it may develop efficient routines for grammatical processing. Languages acquired beyond the critical phase, however, may not be able to automatically and efficiently access these routines, and may have to rely on additional cortical areas outside of, or in the periphery of, primary task-dedicated language areas ((see also Kelly and Garavan (2005)). Interestingly, in contrast to the two late bilingual groups, early bilingual subjects did not show significant differences in functional activity for grammatical processing in the two languages in the areas seen in the late bilinguals, providing further support consistent with the assumption of maturational constraints on efficient grammatical processing. Together, these results suggest that AOA affects the neuronal correlates of grammatical processing, even when subjects have acquired high levels of proficiency. Concurrent acquisition of languages within a critical period for learning grammatical structures cued by word order may lead to more efficient or less resource-demanding grammatical processing than when a second language is learned after the critical period has passed.

A differentiated look at the subject–verb and verb–subject manipulation, furthermore, suggests that not only AOA but also sequence of language acquisition may affect functional activation in grammatical processing. All three groups revealed a higher level of activation on the VS-compared to the SV-sentence structure in the left inferior frontal and inferior parietal cortex, most likely reflecting the higher demands of grammatical processing of the less frequent and noncanonical VS-structure. However, there was a group-by-word order interaction: the L2G-group, for which the VS-sentence structure is marked, showed higher activation for the VS-condition than the L2F group. Although the effect was small, the parameter estimates clearly showed a significant increase in activation of the verb–subject compared to the subject–verb order selectively for the native French speakers who during first language acquisition were less exposed to this structure than subjects in the other two groups who had been frequently exposed to this structure from early on. Further support for maturational influences on the processing of word order comes from a study by Jeong, Sugiura, and Sassa, 2007). In this study, native Chinese and Korean speakers, who had begun to learn English after 12 years on average, were asked to listen either to sentences of their native language (L1) or to English sentences (L2). Based on the fact that both English and Chinese have a basic word order of S–V-O while Korean has an S-O-V word order, the authors predicted greater activation for L2 compared to L1 in the Korean compared to the Chinese group. This selective increase in activation for the Korean group was expected to be localised in areas involved in the processing of word order. As predicted, native Korean compared to native Chinese speakers showed higher activation in several cortical structures including the left IFG when listening to English sentences, compared to sentences in their respective L1. The authors concluded that the observed effect may be caused by the linguistic distance between the language pairs, which is greater between English and Korean than between English and Chinese. Although this is a convincing result which corresponds quite well to our findings, the study's findings are somewhat restricted by the fact that word order was not specifically manipulated. As the authors themselves concede, the observed effects may partially be explained by differences between the (Chinese and Korean) languages other than word order preference.

One might argue that the selective increase in activation for VS structures in the L2G group may simply be an effect of age, since subjects in the L2G group on average were about 12 years older than the subjects in the L2F group (i.e. 41.6 versus 29.3 years of age, respectively). Older age may slow down processing, particularly when processing more complex syntactic structures, which might result in increased activation. However, older and younger adults show similar effects of syntactic complexity on on-line measures (Waters & Caplan, 2001). Apparently, individual speed of processing, not age per se, determines whether syntax-related activation is found in IFG or in other, e.g. parietal, areas (Caplan, 2001).

It is necessary to point out the limitations of our study. The two late bilingual groups had different levels of exposure to L2. According to an important theoretical account, the “Unified Competition Model” (MacWhinney, 2005), efficiency in sentence processing may increase gradually over time as a function of usage and level of competence, even beyond the “critical period”. Thus, in contrast to maturational accounts, the competition model does not predict a pronounced drop in L2 learning abilities beyond this point in time. A valid comparison of the two accounts may only be possible by carefully controlling extent of usage as well as level of competence, and may require longitudinal designs. Our study was not set up to contrast the two accounts. However, if extent L2 usage were able to explain some or all of the variance in L2 processing, one might have expected the late bilingual group with the higher amount of total L2 exposure (i.e. the native French speakers) to show less or no additional activation when processing sentences in L2, compared to L1. This was not the case, as evident from the parameter estimates in Fig. 3.

Our data were acquired in Germany and all participants lived in Germany at the time of scanning. Thus, all participants were exposed predominantly to German at the time of scanning, introducing a “language environment” bias. Ideally, the switch between French and German in L1 and L2 would have been matched by a switch in language environment as well. In addition, the L2G group consisted of native French speakers who had permanently moved to Germany. In contrast, the L2F group mainly consisted of subjects who had spent only a couple of years in France. This resulted in an asymmetry in time of exposure and introduced an additional “time of exposure” bias which is reflected in the proficiency data as well as in the task performance of native German speakers on French items. However, despite a longer exposure to German, we observed the VS-effect in the L2G group. One may speculate that this effect would be even more pronounced if subjects had had less exposure to German. Furthermore, the better task performance for German stimuli in the L2F group should not threaten the selective VS effect of L2G, because the VS effect was computed considering both languages together in each group. Average accuracy (computed across both languages) in the L2F group was 82.8%, and average accuracy in the L2G group was 83.6%, a difference which is not significant ($p = 0.1$).

Finally, we note that the early bilingual group (2L1), despite fulfilling the criteria of bilingualism formally (both languages were learned and practiced before the age of three), revealed a proficiency profile similar to the L2F group, which could partly be explained by the fact that most of the 2L1 subjects were born and had spent most of their lives in Germany. In addition, proficiency data from the 2L1 group showed that although more proficient in German, the L21 group did not reach the proficiency levels of native L1 speakers in both German and French. Thus, although we propose this cautiously, there may be a “proficiency cost” affecting both languages associated with being an early bilingual.

In sum, this study investigated the neural basis of syntactic processing in proficient German–French bilingual subjects. We showed that a fronto-temporal cortico-subcortical network is more involved in word order processing in late compared to early bilingual speakers, suggesting a constraining effect of AOA on syntactic processing. In addition, we were able to demonstrate an effect of sequence of onset of language acquisition on word order processing.

Acknowledgments

This study has been carried out as part of the collaborative research project (SFB 539) funded by the DFG research project “Simultaneous and successive acquisition of bilingualism”, directed by J.M. Meisel. C.B. is supported by the Volkswagen-Stiftung and the BMBF, A.B. was supported by the Deutsche Forschungsgemeinschaft (DFG). We wish to thank two anonymous reviewers for their knowledgeable and constructive comments.

References

Abutalebi, J., Brambati, S. M., Annoni, J. M., Moro, A., Cappa, S. F., & Perani, D. (2007). The neural cost of the auditory perception of language switches: An event-related functional magnetic resonance imaging study in bilinguals. *Journal of Neuroscience*, *27*, 13762–13769.

Badre, D., & Wagner, A. D. (2004). Selection, integration, and conflict monitoring; assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron*, *41*, 473–487.

Bayer, J. (1984). COMP in Bavarian syntax. *The Linguistic Review*, *3*, 209–274.

Behnstedt, P. (1973). *Viens-tu? Est-ce-que tu viens? Tu viens?* Tübingen: Narr.

Bley-Vroman, R. (1989). *What is the logical problem of foreign language learning?* Cambridge: Cambridge University Press.

Caplan, D. (2001). Functional neuroimaging studies of syntactic processing. *Journal of Psycholinguist Research*, *30*, 297–320.

Caplan, D., Alpert, N., & Waters, G. (1999). PET studies of syntactic processing with auditory sentence presentation. *Neuroimage*, *9*, 343–351.

Caplan, D., Alpert, N., Waters, G., & Olivieri, A. (2000). Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Human Brain Mapping*, *9*, 65–71.

Chee, M. W., Tan, E. W., & Thiel, T. (1999). Mandarin and English single word processing studied with functional magnetic resonance imaging. *Journal of Neuroscience*, *19*, 3050–3056.

Chee, M. W., Hon, N., Lee, H. L., & Soon, C. S. (2001). Relative language proficiency modulates BOLD signal change when bilinguals perform semantic judgments. Blood oxygen level dependent. *Neuroimage*, *13*, 1155–1163.

Crinion, J., Turner, R., Grogan, A., Hanakawa, T., Noppeney, U., Devlin, J. T., et al. (2006). Language control in the bilingual brain. *Science*, *312*, 1537–1540.

Crystal, D. (1997). *A dictionary of linguistics and phonetics*. Oxford: Blackwell.

Dapretto, M., & Bookheimer, S. Y. (1999). Form and content: Dissociating syntax and semantics in sentence comprehension. *Neuron*, *24*, 427–432.

De Bleser, R. (2003). The organization of the bilingual lexicon: A PET study. *Journal of Neurolinguistics*, *16*, 439–456.

Dehaene, S., Dupoux, E., Mehler, J., Cohen, L., Paulesu, E., Perani, D., et al. (1997). Anatomical variability in the cortical representation of first and second language. *Neuroreport*, *8*, 3809–3815.

Den Besten, H. (1977). On the interaction of root transformations and lexical deletive rules. *Groninger Arbeiten zur Germanistischen Linguistik*, *20*, 1–78.

Eichheim, H., & Storch, G. (1992). *Mit Erfolg zum Zertifikat* (*Testheft*). Deutschland: Klett Edition.

Embick, D., Marantz, A., Miyashita, Y., O'Neil, W., & Sakai, K. L. (2000). A syntactic specialization for Broca's area. *Proceedings of the National Academic Sciences of the United States of America*, *97*, 6150–6154.

Friederici, A. D. (2004). Processing local transitions versus long-distance syntactic hierarchies. *Trends of Cognitive Science*, *8*, 245–247.

Friederici, A. D., Fiebach, C. J., Schlesewsky, M., Bornkessel, I. D., & von Cramon, D. Y. (2006). Processing linguistic complexity and grammaticality in the left frontal cortex. *Cerebral Cortex*, *16*, 1709–1717.

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J. B., Frith, C. D., & Frackowiak, R. S. (1995). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*(3), 189–210.

Friston, K. J., Ashburner, J., Frith, C. D., Poline, J. B., & Frackowiak, R. S. (1995). Spatial registration and normalization of images. *Human Brain Mapping*, 165–189.

Gabrieli, J. D., Poldrack, R. A., & Desmond, J. E. (1998). The role of left prefrontal cortex in language and memory. *Proceedings of the National Academic Sciences of the United States of America*, *95*, 906–913.

Goethe-Institut (1981). *Prüfungsaufgaben zum Deutschen Sprachdiplom für Ausländer VI 1977–1979*. München: Verlag für Deutsch.

Golestani, N., Alario, F. X., Meriaux, S., Le Bihan, D., Dehaene, S., & Pallier, C. (2006). Syntax production in bilinguals. *Neuropsychologia*, *44*, 1029–1040.

Haller, S., Radue, E. W., Erb, M., Grodd, W., & Kircher, T. (2005). Overt sentence production in event-related fMRI. *Neuropsychologia*, *43*, 807–814.

Hashimoto, R., & Sakai, K. L. (2002). Specialization in the left prefrontal cortex for sentence comprehension. *Neuron*, *35*, 589–597.

Hernandez, A. E., Martinez, A., & Kohnert, K. (2000). In search of the language switch: An fMRI study of picture naming in Spanish–English bilinguals. *Brain Language*, *73*, 421–431.

Hernandez, A. E., Dapretto, M., Mazziotta, J., & Bookheimer, S. (2001). Language switching and language representation in Spanish–English bilinguals: An fMRI study. *Neuroimage*, *14*, 510–520.

Hernandez, A. E., Hofmann, J., & Kotz, S. A. (2007). Age of acquisition modulates neural activity for both regular and irregular syntactic functions. *Neuroimage*, *36*, 912–923.

Hyltenstam, K., & Abrahamsson, N. (2003). *Maturational constraints in second language acquisition*. Oxford: Blackwell.

Illes, J., Francis, W. S., Desmond, J. E., Gabrieli, J. D., Glover, G. H., Poldrack, R., et al. (1999). Convergent cortical representation of semantic processing in bilinguals. *Brain Language*, *70*, 347–363.

Jeong, H., Sugiura, M., & Sassa, Y. (2007). Cross-linguistic influence on brain activation during second language processing: An fMRI study. *Bilingualism: Language and Cognition*, *10*, 175–187.

Juilland, A., Brodin, D., & Davidovitch, C. (1970). *Frequency dictionary of French words*. Paris: Mouton.

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science*, *274*, 114–116.

Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. *Trends of Cognitive Science*, *6*, 350–356.

Kelly, A. M., & Garavan, H. (2005). Human functional neuroimaging of brain changes associated with practice. *Cerebral Cortex*, *15*, 1089–1102.

MacWhinney, B. (2005). A unified model of language acquisition. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 49–67). New York: Oxford University Press.

McDonald, J. (2000). Grammaticality judgements in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics*, *21*, 395–423.

Meisel, J. M. (1989). Early differentiation of languages in bilingual children. In K. Hyltenstam & L. K. Obler (Eds.), *Bilingualism across the lifespan: Aspects of acquisition, maturity, and loss* (pp. 13–39). Cambridge: Cambridge University Press.

Meisel, J. M. (1991). Principles of universal grammar and strategies of language use: On some similarities and differences between first and second language acquisition. In L. Eubank (Ed.), *Point counterpoint: Universal grammar in the second language* (pp. 231–276). Amsterdam, Philadelphia: Benjamins.

Meisel, J. M. (1997). The acquisition of the syntax of negation in French and German: Contrasting first and second language development. *Second Language Research*, *13*, 227–263.

Meschyan, G., & Hernandez, A. E. (2006). Impact of language proficiency and orthographic transparency on bilingual word reading: An fMRI investigation. *Neuroimage*, *29*, 1135–1140.

Moro, A., Tettamanti, M., Perani, D., Donati, C., Cappa, S. F., & Fazio, F. (2001). Syntax and the brain: Disentangling grammar by selective anomalies. *Neuroimage*, *13*, 110–118.

Newman, A. J. (2001). An event-related fMRI study of syntactic and semantic violations. *Journal of Psycholinguistic Research*, *30*, 339–363.

Ni, W., Constable, R. T., Mencl, W. E., Pugh, K. R., Fulbright, R. K., Shaywitz, S. E., et al. (2000). An event-related neuroimaging study distinguishing form and content in sentence processing. *Journal of Cognitive Neuroscience*, *12*, 120–133.

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J. B. (2005). Valid conjunction inference with the minimum statistic. *Neuroimage*, *25*, 653–660.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113.

Perani, D., Dehaene, S., Grassi, F., Cohen, L., Cappa, S. F., Dupoux, E., et al. (1996). Brain processing of native and foreign languages. *Neuroreport*, *7*, 2439–2444.

Perani, D., Paulesu, E., Galles, N. S., Dupoux, E., Dehaene, S., Bettinardi, V., et al. (1998). The bilingual brain. Proficiency and age of acquisition of the second language. *Brain*, *121*(Pt 10), 1841–1852.

Platzack, C. (1983). Germanic word order and the COMP/INFL parameter. *Working Papers in Scandinavian Syntax*, *2*, 1–45.

Roder, B., Stock, O., Neville, H., Bien, S., & Rosler, F. (2002). Brain activation modulated by the comprehension of normal and pseudo-word sentences of different processing demands: A functional magnetic resonance imaging study. *Neuroimage*, *15*, 1003–1014.

Ruoff, A. (1990). *Häufigkeitswörterbuch gesprochener Sprache*. Tübingen: Niemeyer.

Ruschemeyer, S. A., Fiebach, C. J., Kempe, V., & Friederici, A. D. (2005). Processing lexical semantic and syntactic information in first and second language: fMRI evidence from German and Russian. *Human Brain Mapping*, *25*, 266–286.

Schubotz, R. I., & von Cramon, D. Y. (2004). Sequences of abstract nonbiological stimuli share ventral premotor cortex with action observation and imagery. *Journal of Neuroscience*, *24*, 5467–5474.

Tatsuno, Y., & Sakai, K. L. (2005). Language-related activations in the left prefrontal regions are differentially modulated by age, proficiency, and task demands. *Journal of Neuroscience*, *25*, 1637–1644.

van Heuven, W. J., Schriefers, H., Dijkstra, T., & Hagoort, P. (2008). Language conflict in the bilingual brain. *Cerebral Cortex*, doi:10.1093/cercorbhn030

Vingerhoets, G., Van Borsel, J., Tesink, C., van den Noort, M., Deblaere, K., Seurinck, R., et al. (2003). Multilingualism: An fMRI study. *Neuroimage*, *20*, 2181–2196.

Wartenburger, I., Heekeren, H. R., Abutalebi, J., Cappa, S. F., Villringer, A., & Perani, D. (2003). Early setting of grammatical processing in the bilingual brain. *Neuron*, *37*, 159–170.

Wartenburger, I., Heekeren, H. R., Burchert, F., Heinemann, S., De Bleser, R., & Villringer, A. (2004). Neural correlates of syntactic transformations. *Human Brain Mapping*, *22*, 72–81.

Waters, G. S., & Caplan, D. (2001). Age, working memory, and on-line syntactic processing in sentence comprehension. *Psychology and Aging*, *16*, 128–144.

Worsley, K. J., & Friston, K. J. (1995). Analysis of fMRI time-series revisited-again. *Neuroimage*, *2*, 173–181.