Electrophysiological correlates of noun-noun compound processing by non-native speakers of English

Cecile De Cat [1], Ekaterini Klepousniotou [2], Harald Baayen [3]
[1] Linguistics & Phonetics, University of Leeds, UK
c.decat@leeds.ac.uk
[2] Institute for Psychological Sciences, University of Leeds, UK
e.klepousniotou@leeds.ac.uk
[3] Quantitative Linguistics, University of Tübingen, Germany
harald.baayen@uni-tuebingen.de

Abstract

We report on an experimental study of the processing of noun-noun compounds by native and non-native speakers of English, based on Event-Related Potentials recorded during a mask-primed lexical decision task. Analysis was by generalised linear mixed-effect modelling and generalised additive mixed modelling. Non-native processing is found to display headedness effects induced by the mothertongue. The frequency of the constituent nouns and of the intended compounds are also shown to have an effect on processing.

1 Introduction

This study examines the processing of noun-noun compounds by native and non-native speakers of English. Compounds have been extensively studied in the past 40 years from a myriad of viewpoints (Libben and Jarema, 2006; Lieber and Štekauer, 2009). A key concern has been whether the processing of compounds consists in retrieving entities listed in the mind (Butterworth, 1983) or requires decomposition into constituents listed separately (Semenza et al., 1997; Libben, 1998). Dual-routes theories contend that the two processes exist side by side (Sandra, 1990). It is now widely accepted that both constituents are activated during processing, at least in non-lexicalised compounds (Jarema, 2006; Zhang et al., 2012). Noun-noun compounds have also been shown to be processed differently to non-compounds of similar morphological complexity and length, with compounds yielding longer reaction times and different electrophysiological correlates (El Yagoubi et al., 2008).

Endocentric compounds contain a head element (*dust* in (1)) whose lexical category and interpretive features are inherited by the compound and contribute the core of its meaning (e.g. a kind of dust). The other element acts as a modifier of that head.

(1) moon dust ('dust from the moon /dust made of moon /dust with moon-like properties')

Here we focus on endocentric noun-noun compounds (henceforth NNCs), which have been argued to embody an underlying structure (Libben, 2006) that is hierarchical, involving the (possibly recursive) subordination of a modifier to a grammatical head (or a modifier-head compound, as in (2)), with head-directionality that mirrors that of other noun-complement structures in the same language (Zipser, 2013).

(2) [child [amateur [puppet theatre]]]

Proceedings of the First Workshop on Computational Approaches to Compound Analysis, pages 41–52,
Dublin, Ireland, August 24 2014.

Headedness plays a specific role in the processing of NNCs, as demonstrated by research on Italian (which crucially features the two word orders in NNCs). Based on a lexical decision task, El Yagoubi et al. (2008) found priming effects induced by the head, independently of its position in the NNC. Headedeness effects are not distinguishable from position-in-the-string effects in languages such as English. For instance, Jarema et al. (1999) observed no difference in the priming of NNCs by the head or the modifier. Here we take this line of research further, by investigating whether headedness in the mothertongue affects the processing of transparent, irreversible NNCs in highly advanced second language learners of English.

Event-related potentials (ERPs) can provide insight into the neural activity associated with the processing of compounds. Functional interpretations can be inferred from the temporal and spatial characteristics of electromagnetic activity, and ERP components can sometimes reveal the engagement of the cognitive processes involved. Our approach is this paper is exploratory (Otten and Rugg, 2005) and will focus on identifying differences in the amplitude of the EEG signal that can be traced back to properties of the participants (such as their language background) and properties of the compounds (such as their frequency of occurrence, and the frequencies of occurrence of their constituents). Inferences based on previously identified ERP components will be drawn in the discussion as appropriate. Our research questions are: (i) Does non-native processing of NNCs result in different ERP signatures to native processing? (ii) Is non-native processing of NNCs affected by headedness effects from the mothertongue?

2 Materials and methods

We registered the electrophysiological response of the brain to visual stimuli presented in the context of a (masked) primed lexical decision task. Stimuli were irreversible NNCs presented in licit (3-a) and reversed order (3-b).

(3) a. coal dust
b. #dust coal

The participant groups differed in mothertongue: English (control group), Spanish or German (experimental groups). Like English, German features productive compounding, with a head-last structure. Whereas in Spanish, compounds are essentially head-first, and not productive.

2.1 Participants

Ten native British English speakers (4 female, mean age 22;11 years; STD 3;3 years), ten native German learners of English (7 female, mean age 26;5 years; STD 5;7 years) and ten native Spanish learners of English (3 female, mean age 26;11 years; STD 5;3 years) took part in the study. Participants all had initial second-language exposure after 8 years of age, and all scored above 60% on a cloze test from the Cambridge Certificate in Advanced English. All were right-handed based on the Briggs and Nebes inventory (Briggs and Nebes, 1975), had no speech or language difficulties and had normal or corrected-to normal vision.

2.2 Stimuli

Experimental stimuli consisted of prime-target pairs, presented in 4 experimental conditions in a 3 (Group) x 2 (Prime Condition) x 2 (Word Order) design. The prime was either the head (e.g. *dust* in (3)) or the modifier (e.g. *coal* in (3)) of the intended compound.

The Word Order factor had 2 levels: licit (modifier - head, as in (3-a)) or reversed (head - modifier, as in (3-b)). All the NNCs were endocentric and featured a transparent, modification relationship. All items were tested for irreversibility on an independent group of 30 native speakers. The frequency of the licit compounds and their constituent nouns was estimated from the post-1990 data in Google N-grams. To avoid lexicalisation effects, only compounds with very low frequencies were included (i.e. below 3,300 — mean = 359.5, compared with a mean of 279,300 for the constituent nouns).

There was a total of 480 test items (based on 120 compounds), of which 240 are included in the present study (as we focus on the Head Prime condition only). The items were pseudo-randomised into 8 different orders (assigned randomly to participants) and presented in 4 blocks, with a rest in between.

2.3 Procedure

Participants were tested individually in a single session lasting approximately 1.5 hours. Stimuli were presented visually in light grey text on a black background. Each trial began with the visual presentation of a series of exclamation marks (!!!) for 1000 ms, which was a signal for the participant to rest their eyes and blink After a delay of 100 ms a fixation point (+) was presented for 250 ms to signal that the trial was about to begin. After a 100 ms mask (#######), the prime was presented for 100 ms followed by a second mask (for 50 ms) and the target (for 1000 ms). After a delay of 500 ms a question mark (?) appeared for 2000 ms during which time participants had to make a lexical decision about the target (as acceptable or not) by pressing (with their right hand) one of two buttons on a hand-held button box (counterbalanced across participants). Participants were instructed to respond as accurately as possible; accuracy and reaction times (in ms from the onset of the “?") were recorded. After the response (or at the end of 2000 ms if the participant did not respond), there was a delay of 100 ms before the next trial started. The experimental session was preceded by a practice session comprising 20 trials, which was repeated until participants could perform the task and procedure with no errors (usually one or two practice sessions sufficed).

The EEG was recorded (Neuroscan Synamps2) from 60 Ag/AgCl electrodes embedded in a cap based on the extended version of the International 10-20 positioning system (Sharbrough et al., 1991). Additional electrodes were placed on the left and right mastoids. Data were recorded using a central reference electrode placed between Cz and CPz. The ground electrode was positioned between Fz and Fpz. To capture noise articfacts in the EEG signal due to eye movements, electro-oculograms (EOGs) were recorded using electrodes positioned at either side of the eyes, and above and below the left eye. At the beginning of the experiment electrode impedances were below 10 kΩ. The analogue EEG and EOG recordings were amplified (band pass filter 0.1 to 100Hz), and continuously digitised (32-bit) at a sampling frequency of 500 Hz. Data were processed offline using Neuroscan Edit 4.3 software (Compumedics Neuroscan) and filtered (0.1-40Hz, 96 dB/Oct, Butterworth zero phase filter). The effect of eye-blink artifacts was minimised by estimating and correcting their contribution to the EEG using a regression procedure which involves calculating an average blink from 32 blinks for each participant, and removing the contribution of the blink from all other channels on a point-by-point basis. Data were epoched between -100 and 1100 ms relative to the onset of the experimental targets and baseline-corrected by subtracting the mean amplitude over the pre-stimulus interval. Epochs were rejected if participants did not make a response within the allocated time (during presentation of the “?"), or if they made an incorrect response. Subsequently the data was downsampled to 125 Hz. Trial rejection was not done *a priori* but based on the residuals of the modelling, resulting in only 0.7% of discarded data.

3 Results

3.1 Accuracy analysis

The responses on the lexical decision task were analysed with a generalised linear mixed-effect model with a binomial link function, using the `lme4` package, version 1.0-4 (Bates et al., 2013) with the 'bobyqa' optimizer. Only those predictors that contributed to the model fit were retained, as shown in Table 1. The covariate 'Compound Frequency' did not reach significance. The model provided a substantially improved fit compared to the null-hypothesis model with random intercepts for participant and item only.

	Coefficient	Std. Error	Z	p
Intercept	-0.7565	1.7174	-0.4405	0.6596
Word Order: Licit	-0.0828	0.1644	-0.5035	0.6146
L1: German	-0.6339	0.3123	-2.0299	0.0424
L1: Spanish	-0.7670	0.4135	-1.8549	0.0636
Proficiency	3.7191	1.7052	2.1811	0.0292
Word Order: Licit by L1: German	0.8710	0.1474	5.9074	0.0000
Word Order: Licit by L1: Spanish	0.9322	0.1410	6.6101	0.0000

Table 1: Coefficients of a logistic mixed-effects regression model fitted to the accuracy data. The reference level for Word Order is Reversed, and for L1: English

Table 1 indicates that for English speakers, accuracy did not differ for the licit and reversed word order conditions. For non-native speakers, accuracy was higher in the Licit Word Order condition, compared with the Reversed Word Order condition. Across groups, greater proficiency afforded higher accuracy. Figure 1 visualizes this pattern of results.

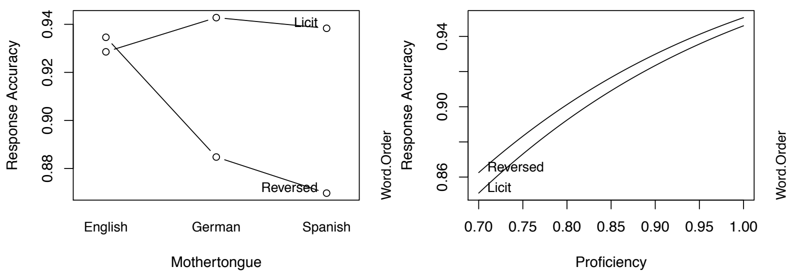

Figure 1: Partial effects of the predictors in the logistic model for response accuracy.

3.2 ERP analysis

We analysed the electrophysiological response elicited by the presentation of compound words with the generalized additive mixed model (GAMM, (Wood, 2006; Tremblay and Baayen, 2010; Baayen, to appear; Baayen et al., in preparation; Kryuchkova et al., 2012)). Generalized additive mixed models extend the generalized linear mixed model with tools (thin plate regression splines, tensor product smooths) for modeling *non-linear* functional relations between one or more predictors and a response variable. GAMMs, as implemented in the mgcv package 1.7-28, offer three important advantages for the analysis of EEG data compared to standard linear models and analysis of variance. First, GAMMs are optimized for dealing with non-linear functional relations between a response (here, the amplitude) and one or more numerical predictors (resulting in wiggly curves, wiggly surfaces, or, in the case of more than two predictors, wiggly hypersurfaces). Second, GAMMs decompose the EEG amplitude into a sequence of additive components, thereby affording the analyst a toolkit for separating out partial effects due to different kinds of predictors (e.g., language group, time, compound frequency, constituent frequency). Third, GAMMs can capture AR1 autocorrelative processes in the signal, and therefore protect against anti-conservative p-values and mistakingly taking noise for complex EPR signatures (as has been shown to occur by Tanner et al., 2013).

We include for analysis only trials that elicited a correct response. The time window analysed was limited to 0–800 ms, time-locked to the onset of stimulus presentation. Autocorrelations in the residual error were removed by including in the GAMM an autocorrelation parameter $\rho = 0.9$ for AR1 error for each basic time series in the data (the time series amplitudes for each unique combination of subject and item). Inclusion of ρ was essential for removing most of the autocorrelational structure from the model's residuals.

A. parametric coefficients	Estimate	Std. Error	t-value	p-value
Intercept (English Reversed)	-0.6815	1.8695	-0.3645	0.7155
Compound Frequency (English Reversed)	0.0659	0.0756	0.8711	0.3837
English:Licit	1.0720	0.1845	5.8103	< 0.0001
German:Reversed	0.6172	2.5967	0.2377	0.8121
German:Licit	0.8199	2.5977	0.3156	0.7523
Spanish:Reversed	0.0311	2.5986	0.0120	0.9905
Spanish:Licit	-3.6624	2.6002	-1.4085	0.1590
Comp. Frequency:English Licit	-0.2747	0.0392	-7.0097	< 0.0001
Comp. Frequency:German Reversed	-0.0577	0.0405	-1.4254	0.1540
Comp. Frequency:German Licit	-0.0826	0.0397	-2.0837	0.0372
Comp. Frequency:Spanish Reversed	-0.1139	0.0414	-2.7536	0.0059
Comp. Frequency:Spanish Licit	0.2361	0.0404	5.8473	< 0.0001
B. smooth terms	edf	Ref.df	F-value	p-value
smooth in Time English:Licit	8.5809	8.7860	11.5648	< 0.0001
diff. curve Time: German:Licit	1.0111	1.0212	0.1285	0.7255
diff. curve Time: Spanish:Licit	6.7504	7.8925	4.4964	< 0.0001
diff. curve Time: English:Reversed	1.9025	2.3906	1.0696	0.3436
diff. curve Time: German:Reversed	1.0074	1.0141	0.4174	0.5210
diff. curve Time: Spanish:Reversed	1.0069	1.0095	1.6952	0.1925
tensor product surface F1 and F2 (English, Licit)	3.0189	3.0349	2.1154	0.0951
diff. surface German:Licit	11.2569	12.3579	7.2697	< 0.0001
diff. surface Spanish:Licit	12.9312	13.6137	60.1585	< 0.0001
diff. surface English:Reversed	3.9839	4.0083	17.6082	< 0.0001
diff. surface German:Reversed	9.0655	10.4566	5.5875	< 0.0001
diff. surface Spanish:Reversed	14.7736	14.9639	28.2189	< 0.0001
random intercepts Compound	107.6142	111.0000	34.7869	< 0.0001
by-subject random wiggly curves Trial	163.4484	267.0000	43.8796	< 0.0001
by-subject random wiggly curves Time	170.5793	267.0000	2.4442	< 0.0001

Table 2: Generalized additive mixed model fitted to the amplitude of the electrophysiological response of the brain to English compounds at channel C1.

In what follows, we focus on channel C1, which revealed a pattern of results typical for surrounding channels. The amplitude of the EEG signal was modeled (without any prior averaging) as an additive function of Word order (Licit vs. Reversed), Compound Frequency, the Constituent Frequency of Modifier and of Head, and Participant Group (English, German, Spanish). Proficiency did not reach significance and did not improve the model fit significantly, so we did not include this predictor in the final model.

GAMMs currently can only accomodate interactions of smooths with a single factor. In order to study the interaction of speaker group and word order, we therefore created a new factor GO with

as levels English:Licit, English:Reversed, German:Licit, German:Reversed, Spanish:Licit, and Spanish:Reversed, using treatment contrasts with as reference level English:Reversed. In the parametric part of the model (the upper half of Table 2), the coefficients for the main effect of GO and its interaction with compound frequency are to be interpreted in the familiar way, with the interaction terms specifying differences in the slope of compound frequency for the non-reference levels of GO.
GO also interacted with the constituent frequencies. For this three-way interaction, we recoded GO as an ordered factor, which is how the `bam` function of the MGCV package is instructed to construct a reference surface (in our implementation, for English:Licit) and difference surfaces for the other factor levels with respect to the standard compound forms as read by English native speakers.
Table 2 summarizes the GAMM fitted to the amplitude of the EEG signal at channel C1. First consider the parametric part of the model, presented in the upper half of the table, which concerns the main effect of GO and its interaction with log-transformed compound frequency. This interaction is summarized in Figure 2. Black lines denote the Licit Word Order condition, grey lines the Reversed Word Order condition. Compound frequency did not have much of an effect in the Reversed conditions.

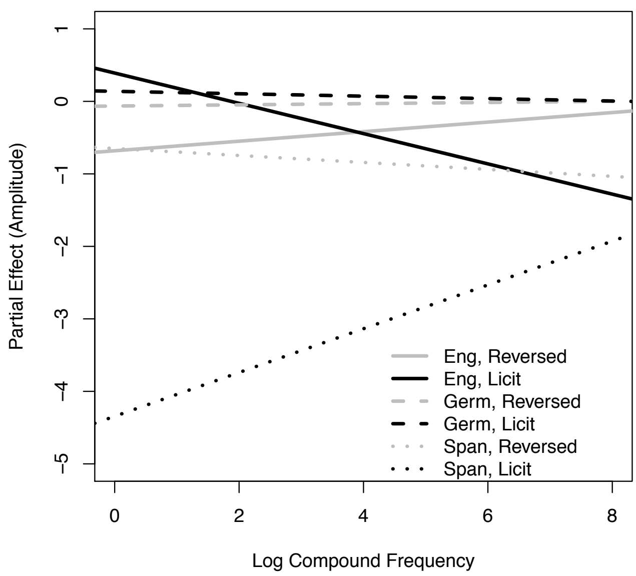

Figure 2: The three-way interaction of Participant Group, Grammaticality, and Compound Frequency.

For English (solid lines), a compound frequency is present in the licit condition, with greater compound frequencies inducing more negative amplitudes. For German (dashed lines), the slope was close to zero in both conditions, indicating the absence of a frequency effect. The Spanish speakers (dotted lines) revealed a regression line with an opposite slope to that for the English speakers in the Licit condition, and with a much lower intercept. This reversal of the slope, as compared to English, may be a consequence of the fact that in Spanish, translation equivalents would be expressed with the opposite constituent order.
The non-parametric part of the model, reported in the lower half of Table 2, handles non-linear effects in the model, using thin plate regression splines for wiggly curves and tensor product smooths for wiggly surfaces. The first row of the non-parametric subtable summarizes a smooth in time for English

licit compounds. This smooth is visualized in the left panel of Figure 3, together with its 95% confidence interval. The model required 8.78 effective degrees of freedom (edf) to capture a (significant) positive inflection around 300 ms post stimulus onset. (Higher edfs indicate greater wiggliness.) The next 5 rows in Table 2 describe the difference curves for the remaining levels of GO. The only level for which this difference curve is significant is Spanish:Licit. The second panel of Figure 3 presents this difference curve, which required 7.89 effective degrees of freedom. As the difference curve is significantly above the X-axis around 300 ms post stimulus onset, and significantly below the X-axis after 600 ms, we conclude that the Spanish speakers reading licit compounds had a higher positivity around 300 ms compared to the English speakers reading the same compound, combined with more negative amplitudes after 600 ms post stimulus.

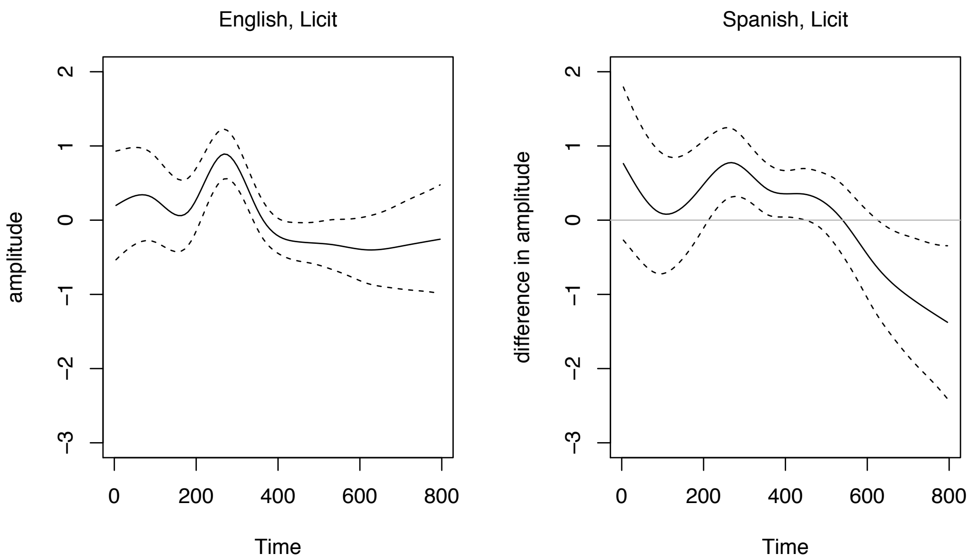

Figure 3: The interaction of participant Group, Grammaticality, and Time. The left panel shows the smooth for English in the Licit Word Order condition; the right panel shows the difference curve with respect to the left panel for the Spanish participants.

EEG amplitudes were also modulated by an interaction of the constituent frequencies by GO, which we modeled with a tensor surface for English:Licit and difference tensor surfaces for the other levels of GO. The second set of 6 rows in Table 2 present the summary statistics, and Figure 3 the smoothed surfaces. The upper left panel presents the reference smooth for English native speakers reading compounds in their licit order. For channel C1, this surface is not well-supported statistically ($p = 0.095$), but at neighboring channels (e.g., Cz, FC1) higher-frequency constituents elicited significantly higher amplitudes. Interestingly, when the constituents are reversed, significantly more negative amplitudes for compounds with high constituent frequencies are observed for native English speakers, as shown in the lower left panel. German speakers show a similar pattern with more negative amplitudes for both licit and reversed compounds (center panels). The strongest negativities are present for Spanish speakers in the licit condition (upper right). In the reversed condition, Spanish speakers show a pattern of somewhat increased negativity (lower right) that, however, does not vary much with constituent frequency.

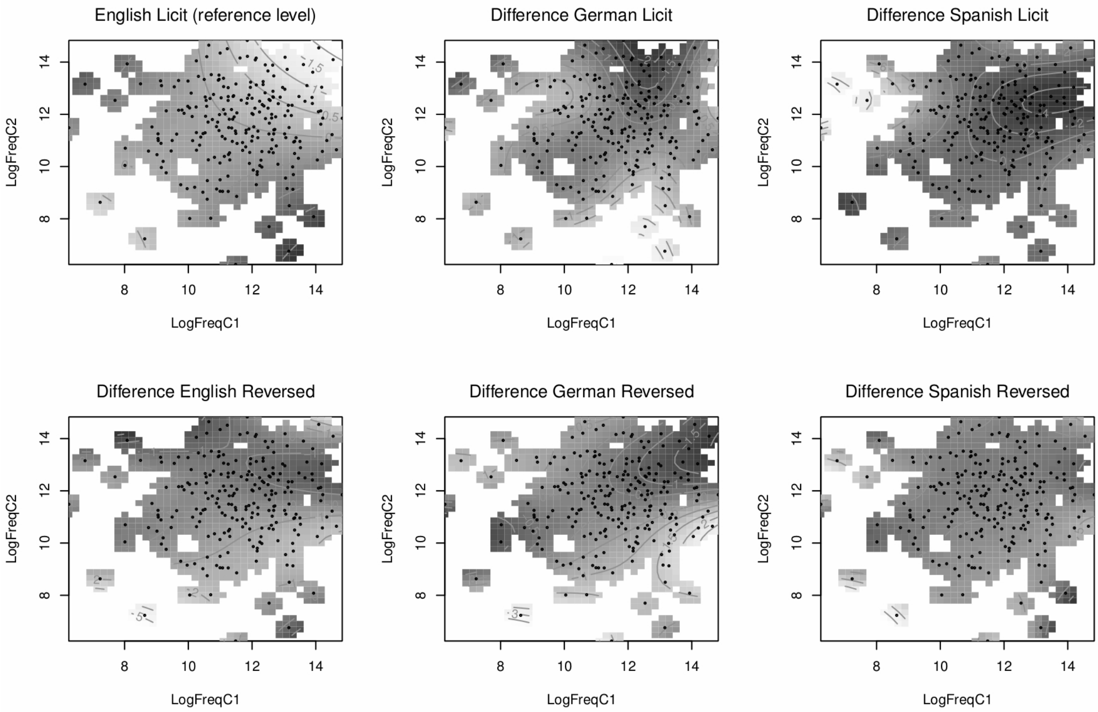

Figure 4: The interaction of left and right constituent frequency by grammaticality and language group. The upper left panel presents the smooth surface for English:Licit, the remaining panels present difference surfaces with respect to the English:Licit condition. Darker shades of gray indicate more negative amplitudes. Contour lines are 0.5 units apart in panels 1, 2, 5, and 6; they are 2 units apart in panel 3, and 1 unit apart in panel 4.

The final three rows of Table 2 specify the random-effects structure of the model. Random intercepts for compound were included in order to allow for differences in baseline amplitude across compounds. For subjects, two random wiggly curves were included. The first models changes in amplitude as subjects go through the experiment. The second models subject-specific changes over within-trial time. The random wiggly curves are the nonlinear equivalent of what in the context of a linear mixed-effects model would be 'random straight lines' obtained by combining random intercepts with random slopes. For EEG data, where amplitude changes non-linearly with time, the flexibility of penalized and shrunk regression splines is essential.

4 Discussion

The non-native participants performed the lexical decision task with a high level of accuracy. For the licit compounds, accuracy was comparable to that of native speakers. For reversed compounds, accuracy dropped slightly, from around 94% to around 88%. From this, we conclude, first, that all subjects have acquired NNC structures in English, and second, that non-native speakers are more likely to accept novel noun combinations as English compounds.

Knowledge of whether a two-word combination is in fact licit in English can arise from two sources. On the one hand, speakers may be familiar with the compound, as evidenced by an effect of compound frequency. For the native English speakers resonding to licit compounds, an effect of compound frequency was indeed present in the EEG amplitudes. On the other hand, speakers may infer the intended meaning from the constituents (e.g., English *beach ball* indexing German *Wasserball*, 'water ball'). Constituent effects were well attested in the EEG amplitudes. Interestingly, for English speakers, constituent frequency effects gave rise to more positive amplitudes in the licit condition (significantly at neighboring channels) whereas in the reversed condition, amplitudes were more negative for higher-frequency constituents. In other words, when English speakers are confronted with reversed compounds, which for them are actually novel compounds, the compound frequency effect disappears, and a constituent frequency effect emerges that is opposite in sign to that for normal compounds.

Of the non-native speakers of English, only the Spanish speakers revealed a compound frequency effect, with a slope opposite in sign to that for the English speakers. If higher amplitude in the signal indicates increased processing effort, the effect of the frequency of the intended compound could be interpreted as facilitating in the Licit Word Order condition in the native speakers but inhibiting in the Spanish group (and without much effect in the German group). We hypothesize that Spanish speakers find licit English compounds more difficult precisely because in their native language, the order of the constituents would have been reversed. It is only these speakers that have a word order conflict to resolve.

All speakers (non-native as well as English) responding to reversed (i.e., for them, novel) compounds, show more negative amplitudes for compounds with higher constituent frequencies. We interpret this as evidence for constituent-driven, decompositional processing. The especially pronounced negativities for Spanish speakers in the Licit Word Order context (which go hand in hand with a positive slope for compound frequency) suggest that for these speakers increased processing resources are called upon to resolve the conflict between English and Spanish constituent order, in spite of native-like performance in the evaluation of compounds in that condition.

A positive peak around 300 ms post-stimulus was found in all groups in both conditions, and exacerbated in the Spanish group in the Licit Word Order condition. This peak could be interpreted as a P300, indexing attentional resources. El Yagoubi et al. (2008) found that right-headed NNCs in Italian yielded a greater P300 and interpreted this as evidence that processing this marked (but in Italian equally grammatical) word order required increased attentional resources. If the P300 observed here reflects a peak of attentional engagement, we expect its amplitude to predict scores on an Attention Network Task (Fan et al., 2005) — something we will investigate in the next phase of this study.

With respect to the absence of a significant N400 effect between the Word Order conditions, we first note that the N400 may vanish due to familiarization, and also to masked priming (Coulson et al., 2005; Brown and Hagoort, 1993). However, and perhaps more importantly, reversed compounds are not semantically anomalous. To the contrary, they invite interpretation and, as we have documented, give rise to constituent-driven processes of interpretation. From this perspective, an N400 would then characterize the processing of semantic anomalies that cannot be resolved through morphological processing.

5 Concluding remarks

This study set out to investigate (i) whether non-native processing of NNCs results in different ERP signatures compared to native processing, and (ii) whether non-native processing of NNCs is affected by constituent order in the mothertongue. Analysis of the EEG amplitudes revealed that English native speakers read licit compounds using both whole-word information (as indexed by compound frequency) in congruence with constituent information (as indexed by constituent frequency with a positive effect) whereas non-native speakers and English speakers reading novel (reversed) compounds resort to decompositional interpretation indexed by a negative effect on amplitudes. Further-more, Spanish readers undergo interference from the different constituent order possibilities in their own language, leading to a reversed compound frequency effect and strongly enhanced constituent frequency effects (with a negative sign) when reading English licit compounds.

This pattern of results is, for native speakers, consistent with the early effects of compound frequency observed using eye-tracking by, e.g., Kuperman et al. (2008, 2009) and Miwa et al. (2014) for English, Finnish, and Japanese respectively. The importance of constituent-driven processing for non-native speakers is reminiscent of the decompositional eye-movement patterns of less-proficient readers reported by Kuperman & Van Dyke (2011).

We conclude with noting that the insights gleaned from the EEG amplitudes would not have been possible without generalized additive mixed models. At the same time, we believe we are only seeing the tip of the iceberg. For instance, the model can be improved by allowing the interaction of the constituent frequencies by group and constituent order to vary with time, using a five-way tensor product smooth. Two considerations have withheld us from following up on this considerably more complex model. First, without specific hypotheses as a guide, interpretation becomes extremely difficult. Second, we are concerned that with a relative small number of compounds (120), overfitting might become an issue. For future research specifically addressing the development over time of constituent (and whole-word) frequency effects, we recommend designs with larger numbers of compounds.

6 Acknowledgments

This project was financed by pump-priming funds from the University of Leeds' Faculty of Arts and by a British Academy Quantitative Skills Acquisition award (SQ120066) to the first author. Many thanks to Antoine Tremblay for help with the initial data preparation, to Cyrus Shaoul for friendly technical and coding advice, to Jacolien van Rij for helpful suggestions for the gamm analysis and to Raphael Morschett, Chris Norton, Kremena Koleva and Natasha Rust for the data collection and pre-processing.

References

R.Harald Baayen, Jacolien van Rij, Cécile De Cat, and Simon Wood. in preparation. Autocorrelated errors in experimental data in the language sciences: Some solutions offered by generalized additive mixed models.

R. Harald Baayen. to appear. *Analyzing Linguistic Data. A Practical Introduction to Statistics Using R (second, augmented edition).* CUP, Cambridge.

Douglas Bates, Martin Maechler, Ben Bolker, and Steven Walker, 2013. *lme4: Linear mixed-effects models using Eigen and S4.* R package version 1.0-4.

G.G. Briggs and R.D. Nebes. 1975. Patterns of hand preference in a student population. *Cortex*, 11:230–238.

Colin Brown and Peter Hagoort. 1993. The processing nature of the n400: Evidence from masked priming. *Journal of Cognitive Neuroscience*, 5(1):34–44.

B. Butterworth. 1983. Lexical representation. In B. Butterworth, editor, *Language Production*, pages 257–294. Academic Press, San Diego, CA.

S. Coulson, Kara D. Federmeier, C. Van Petten, and Marta Kutas. 2005. Right hemisphere sensitivity to word- and sentence-level context: Evidence from event-related brain potentials. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 31:129–147.

Radouane El Yagoubi, Valentina Chiarelli, Sara Mondini, Gelsomina Perrone, Morena Danieli, and Carlo Semenza. 2008. Neural correlates of Italian nominal compounds and potential impacts of headedness effect: An ERP study. *Cognitive Neuropsychology*, 25(4):559–581.

Jin Fan, Bruce D. McCandliss, John Fossella, Jonathan I. Flombaum, and Michael Posner. 2005. The activation of attentional networks. *NeuroImage*, 26(2):471–479.

Gonia Jarema, C. Busson, R. Nikolova, K. Tsapkini, and Gary Libben. 1999. Processing compounds: A cross-linguistic study. *Brain and Language*, 68:362–369.

Gonia Jarema. 2006. Compound representation and processing: A cross-language perspective. In Gary Libben and Gonia Jarema, editors, *The Representation and Processing of Compound Words*, pages 45–70. OUP, Oxford.

T. Kryuchkova, B. V. Tucker, L. Wurm, and R. H. Baayen. 2012. Danger and usefulness in auditory lexical processing: evidence from electroencephalography. *Brain and Language*, 122:81–91.

V. Kuperman and J.A. Van Dyke. 2011. Effects of individual differences in verbal skills on eye-movement patterns duing sentence reading. *Journal of memory and language*, 65(1):42–73.

V. Kuperman, R. Bertram, and R. H. Baayen. 2008. Morphological dynamics in compound processing. *Language and Cognitive Processes*, 23:1089–1132.

V. Kuperman, R. Schreuder, R. Bertram, and R. H. Baayen. 2009. Reading of multimorphemic Dutch compounds: Towards a multiple route model of lexical processing. *Journal of Experimental Psychology: HPP*, 35:876–895.

Gary Libben and Gonia Jarema, editors. 2006. *The Representation and Processing of Compound Words*. OUP, Oxford.

Gary Libben. 1998. Semantic transparency in the processing of compounds. *Brain and Language*, 61:30–44.

Gary Libben. 2006. Why study compound processing? An overview of the issues. In Gary Libben and Gonia Jarema, editors, *The Representation and Processing of Compound Words*, pages 1–22. OUP, Oxford.

Rochelle Lieber and Pavol Štekauer, editors. 2009. *The Oxford Handbook of Compounding*. Oxford University Press, Oxford.

Koji Miwa, Gary Libben, Ton Dijkstra, and Harald Baayen. 2014. The time-course of lexical activation in japanese morphographic word recognitin: Evidence for a character-driven processing model. *The Quarterly Journal of Experimental Psychology*, 67(1):79–113.

Leun Otten and Michael Rugg. 2005. Interpreting event-related brain potentials. In Todd Handy, editor, *Event-related potentials: A methods handbook*, pages 3–17. MIT Press, Cambridge, MA.

D. Sandra. 1990. On the representation and processing of compound words: Automatic access to constituent morphemes does not occur. *The Quarterly Journal of Experimental Psychology*, 42A:529–567.

Carlo Semenza, C. Luzzatti, and S. Carabelli. 1997. Morphological represntation of compound nouns: A study on Italian aphasic patients. *Journal of Neurolinguistics*, 10:33–43.

F. Sharbrough, G.E. Chatrian, R.P. Lesser, H. Luders, M. Nuwer, and T.W. Picton. 1991. American electroencephalographic society guidelines for standard electrode position nomenclature. *Journal of Clinical Neurophysiology*, 8:200–202.

Darren Tanner, Kayo Inoue, and Lee Osterhout. 2013. Brain-based individual differences in online L2 grammatical comprehension. *Bilingualism: Language and Cognition*, 17(2):277–293.

Antoine Tremblay and R. Harald Baayen. 2010. Holistic processing of regular four-word sequences: A behavioral and ERP study of the effects of structure, frequency, and probability on immediate free recall. In D. Wood, editor, *Perspectives on formulaic language: Acquisition and communication*, pages 151–173. The Continuum International Publishing Group., London.

Simon Wood. 2006. *Generalised additive models: An introduction with R*. Chapman and Hall/CRC, Boca Raton, FL.

J. I. E. Zhang, Richard C. Anderson, Qiuying Wang, Jerome Packard, Xinchun Wu, Shan Tang, and Xiaoling Ke. 2012. Insight into the structure of compound words among speakers of chinese and english. *Applied Psycholinguistics*, 33(4):753–779.

Katharina Zipser. 2013. Proto-language, phrase structure and nominal compounds. Which of them fit together? In *Poster presented at ICL 2013, Geneva*.