Running head: Impaired syntactic prediction in SLI

IMPAIRED SYNTACTIC PREDICTIONS
IN CHILDREN WITH SPECIFIC LANGUAGE IMPAIRMENT

Arild Hestvik[1], Baila Epstein[2], Valerie Shafer[3] & Richard G. Schwartz[3]

[1]Department of Linguistics and Cognitive Science, University of Delaware

[2]Speech Communication Arts & Sciences, Brooklyn College, CUNY

[2]PhD Program in Speech and Hearing Sciences, CUNY Graduate Center

Corresponding author: Arild Hestvik, Department of Linguistics and Cognitive Science, University of Delaware, 42 East Delaware Ave, Newark, DE. E-mail: arild@udel.edu.
Telephone: 302-831-2296. Fax: 302-831-6896.

1. Introduction

Specific Language Impairment (SLI) is a childhood language development impairment that affects about 7% of children, with greater prevalence in boys than girls, and occurs in the absence of obvious neurological or socio-behavioral problems. Furthermore, SLI children have normal hearing and intelligence (Leonard, 1998). A characteristic of SLI (perhaps limited to a subgroup, cf. Friedmann and Novogrodsky (2008)) is a significant syntactic impairment in the mastery of constructions involving the "Displacement operation", e.g. Wh-questions (Deevy and Leonard, 2004; Marinis and Van der Lely, 2007) and relative clauses (Friedmann and Novogrodsky, 2004; Novogrodsky and Friedmann, 2006; Friedmann and Novogrodsky, 2007; Friedmann and Novogrodsky, 2008; Zhang and Zhang, 2008; Friedmann and Novogrodsky, 2011).One approach to SLI is that it is a reflection of a deficit in the grammatical knowledge system (Clahsen, 1989; Clahsen, 2008), perhaps due to deficits in the innately specified constraints on grammar induction. The alternative hypothesis is that SLI children have the same initial predispositions as their typically developing peers, but are impeded in their language learning by non-linguistic deficits in perception or information processing (Joanisse and Seidenberg, 1998; Leonard, 1998; Joanisse and Seidenberg, 2003). Under the first, "knowledge deficit view", the impairment is statically linked to an atypical linguistic knowledge state, and entails that SLI children have a normal processing system but an atypical grammatical system. Under the second, "processing deficit" view, SLI children have an atypical processing system, but have the same grammatical system as typically developing children. A third logical possibility is that SLI children have *both* a processing deficit *and* a knowledge deficit, where the knowledge deficit arises from a processing deficit. This model is harder to test because competence can only be inferred from performance data. If SLI children have both a processing deficit and a knowledge deficit, it is harder to tease apart from the performance data which level the effect comes from. To simplify the problem, we here only consider a processing model where SLI children have attained the normal grammar, but a processing deficit impairs them in putting this intact knowledge to use. This dichotomy can be viewed as saying that the human language system consists of a competence vs. a performance system, and that language impairments can in principle affect either the competence system or the performance system.

In the current study, we examined SLI children's real-time comprehension of filler-gap dependencies in relative clauses, with the aim of determining whether previously observed lack of immediate gap-filling after the verb in relative clauses, is due to a competence deficit, or a performance deficit. Below, we first give a brief overview of filler-gap dependencies, review previous behavioral findings, and then develop the predictions for differentiating a competence deficit from a performance deficit.

1.1 Theoretical background

Filler-gap dependencies are created by the fundamental syntactic operation of *displacement* in natural language; a structural permutation of basic syntactic elements, with specific syntactic and semantic consequences. For example, object relative clauses are widely viewed as arising from the permutation of the object of a verb to a noun phrase-initial position, creating a structure that allows a clausal expression to function as a restrictive modifier of that noun phrase. To illustrate, the structure in (1b) is derived from the clause in (1a) by permutation of the object to the initial position of the noun phrase (Kayne, 1994):

(1) a. [$_S$ the hippo [$_{VP}$ kissed [$_{NP}$ the zebra] [$_{PP}$ on the nose]]]

b. [$_{NP}$ **[$_{NP}$ <u>the zebra</u>]** [that [$_S$ the hippo [$_{VP}$ kissed **[$_{NP}$ <u>*e*</u>]** [$_{PP}$ on the nose]]]

In generative linguistic theory, the permutation of the noun phrase leaves behind a gap (denoted *e*) in (1b), which represents the underlying grammatical function of the displaced noun phrase. During comprehension of such sentences, the parsing system must first identify *the zebra* as a displaced element, and next associate it with the abstract, phonetically silent gap position that it originated from. The real-time cognitive process of computing where displaced phrases come from is known in the psycholinguistic literature as the “gap-filling” process. Gap-filling in relative clauses is required by the listener to compute the grammatical and semantic functions of the relativized noun phrase and the role it plays in the semantics of the modifying clause: both as the head of its own noun phrase, and as the argument of a predicate that restricts the interpretation of that noun phrase. This computation must take place nearly instantaneously, and the “active filler hypothesis” (Crain and Fodor, 1985; Frazier and Flores d'Arcais, 1989) holds that the language processing system achieves this by constantly making predictions about where in the structure the gap is most likely to be located.

1.2 Review of previous behavioral studies of gap-filling in SLI

Children with SLI have difficulties producing and comprehending sentences containing filler-gap dependencies in relative clauses (Schuele and Nicholls, 2000; Schuele and Tolbert, 2001; Stavrakaki, 2001; Friedmann and Novogrodsky, 2004; Novogrodsky and Friedmann, 2006) as well as in Wh-questions (Deevy and Leonard, 2004). According to a proposal in the knowledge deficit category (van der Lely, 1998; Van der Lely and Battell, 2003; van der Lely, in press), the core characteristic of SLI is that they have a deficit in their grammatical knowledge about how to represent grammatical dependencies. Specifically, van der Lely suggests that SLI children only "optionally" represent a gap in Wh-questions and relative clauses. This would mean that some of the time they represent the string in (1b) as in (2):

(2) $[_{NP}$ **$[_{NP}$ the zebra]** [that $[_{S}$ the hippo $[_{VP}$ kissed $[_{PP}$ on the nose]]]

Using this representation, with no link between the relativized noun and object position of the verb, *the zebra* will not be integrated into the meaning of the verb *kissed* and the correct comprehension of the entire string cannot be computed. This representational error would predict comprehension and production difficulties. However, a clear consequence of putting the deficit at the level of the competence grammar is that there should be no interaction with processing factors. I.e., one would not expect a grammatical deficit to be modulated by processing demands. However, as discovered by (Deevy and Leonard, 2004), the severity of SLI children's comprehension problems with Wh-questions correlate with the increased processing demands posed by long Wh-questions (e.g. "Who did Bill say that Mary liked?") compared to short Wh-questions ("Who did Bill like?").

If SLI is caused by a processing deficit, one would look outside the grammatical system for an explanation for why filler-gap dependencies are vulnerable in SLI children, for example, by linking it to working memory limitations (Montgomery, 1995; Weismer, 1996; Weismer et al., 1999; Montgomery, 2006). Arguably, long Wh-questions pose greater processing demands because the filler must be kept in working memory for a long time and over a greater structural distance. (Deevy and Leonard, 2004)'s finding lends support to a processing account of gap-filling problems in SLI, and cannot easily be predicted by a static knowledge deficit theory.

One short-coming of the methodology in (Deevy and Leonard, 2004) is that it examined comprehension with an off-line, picture pointing task. This task measures the end-point of the sentence comprehension process, and does not directly reveal the nature of the representation that feeds into the comprehension system. In order to answer the question of whether SLI children's interpretation is generally based on structures with gaps like (1b) or without gaps like (2), their processing must be probed by methods that examine the real-time, moment-by-moment incremental structure building during processing of filler-gap sentences. In the adult and typically developing child literature, real-time representation of gaps has been assessed via measures of *reactivation* at the gap site of lexical or semantic information associated with the filler. For example, (Love and Swinney, 1997; Love, 2007), using a cross-modal picture priming task, demonstrated that young children (mean age 5;6) exhibited increased semantic activation at the gap site of the appropriate filler in sentences like (3)

(3) The zebra that ^{1}the hippo had kissed^{2} on the nose, ran far away.

In the category decision task employed by (Love, 2007), picture probes appear at either temporal position 1 or 2, and children make picture decisions for pictures either primed by the filler or not. A priming effect was observed for pictures of the relativized nouns, here "zebra", at position 2, but not at position 1, which is interpreted to mean that the children immediately reactivate the filler at this position—i.e., as a consequence of constructing a gap there. This finding was replicated by (Roberts et al., 2007) for gaps in double object constructions like *John saw the peacock to which the ^{1}small penguin gave the nice birthday present^{2} in the garden last weekend* (although only typically developing children with high verbal memory span exhibited the effect, which could be due to the complexity of their stimuli sentences).

This method has also recently been used to examine whether SLI children construct gaps at the same temporal positions as their typically developing peers, or at all. (Marinis and Van der Lely, 2007) used the cross-modal picture priming method to test whether SLI children exhibited reactivation in sentences like *Who did Balloo give[1] the long[2] carrot to [3][e] at the farm?* They found that whereas the typically developing control children exhibited priming at position 3, SLI children did not. In fact, SLI children showed priming at position 1, immediately following the verb, which the authors interpret as showing that SLI children rely on lexical thematic information to compute filler-gap dependencies. One problem, however, with using cross-modal picture priming tasks with SLI children is that the dual, or divided attention nature of the task could cause a confound: SLI children generally have more limited verbal working memory resources that their typically developing peers (Weismer et al., 1999; Montgomery, 2000). The effect in cross-modal lexical priming studies depends on subjects paying full attention and processing the stimuli in both modalities at the same time. If SLI children have reduced working memory resources, this could make it difficult for them to perform the task, which directly would result in lack of reactivation effects for task demand reasons, rather than underlying grammatical reasons. To address this problem, Hestvik et al. (2010) used a much less demanding unary picture-naming task instead of a binary picture categorization task, in a study examining the same constructions as Love and Swinney (1997). The paradigm was identical to the paradigm in Love (2007), except children simply make speeded naming of the pictures they see, with naming latency as the dependent measure. Using the same materials as in (Love, 2007), picture probes were presented at positions 1 and 2 in sentences like *The zebra that the hippo on the hill[1] had* *kissed*[2] *on the nose ran far away*. (Control position 1 was purposefully moved further away from the relative clause noun to attenuate repetition priming effects.) The finding was that the SLI children showed (repetition) priming effects at position 1 and no significant reactivation priming at the gap position 2, whereas the opposite was true for the age-matched group of control children. This finding of Hestvik et al. (2010) are similar to that of Marinis and van der Lely (2007) in not observing priming at the gap position, although the latter study observed priming at the verb for the SLI children. Marinis and Van der Lely (2007) speculates that this shows that SLI children interpret Wh-questions (and presuambly by extension relativized nouns) via lexical-thematic information.

In sum, all the behavioral studies reviewed above provide evidence that SLI children do not reactivate fillers at gap positions in Wh-questions and relative clauses. However, the studies leave a question unanswered: Could it be that SLI children merely differ from typically developing children in the latency with which they perform gap-filling? This hypothesis is motivated by the finding that SLI children exhibit generally slower processing on a range of linguistic and non-linguistic tasks (Kail, 1994; Windsor and Hwang, 1999; Lahey et al., 2001; Miller et al., 2006; Leonard et al., 2007). If SLI children are merely temporally delayed in identifying the filler-gap relationship during sentence processing, it would follow that antecedent reactivation should not be observed at the immediate temporal offset of the verb in sentences like (1b/3), but it could be observed at a temporally later position, depending on the extent of the processing delay. This hypothetical “delayed reactivation” would not be detectable by the behavioral experimental techniques reviewed above: The cross-modal priming experiments reviewed only sample reactivation at the single “correct” position, and will therefore be unable to observe delayed gap-filling. This could be remedied by testing multiple downstream positions, but this again is made difficult by the fact that generalized slowing doesn't predict exactly *when* SLI children might fill the gap, however delayed. "Downstream" sampling therefore runs the risk of resource-intensive shots in the dark.

In order to overcome the methodological limitations of cross-modal priming techniques, and measure whether SLI children exhibit temporally delayed gap-filling, we employed the "filled gaps" event-related brain potentials paradigm developed in Hestvik et al (2007). In this paradigm, the parser's predictions and expectations about where a gap should occur is experimentally violated, and the latency of the brain response to the violation reveals how fast the violation is detected. This technique allows us to directly measure whether SLI children make any gap-predictions at all, and if they do, whether the latency is similar or delayed in comparison to typically developing children.

1.1 The current study: ERP responses elicited by expectancy violations

Before we explain our approach to addressing this question, consider first the details of the gap-filling process. When listeners first encounter a filler, they know that a phrase will be missing somewhere else in the sentence—i.e., the sentence will contain a gap further downstream. For example, after parsing "The zebra that the hippo…", the parser will anticipate a transitive verb and pro-actively predict that the gap should occur after this verb—this is the "active filler strategy" (Stowe, 1986; Frazier and Flores d'Arcais, 1989; Stowe et al., 1991; Gibson and Hickok, 1993; Frazier and Clifton, 1995). Once a transitive verb is encountered and a gap is posited after the verb, the parser then next expects that subsequent categorical structure is consistent with having a gap after the verb. This prediction would be met, for example, by a non-argument prepositional phrase, as in "The zebra that the hippo kissed *on the nose*." The preposition confirms the gap prediction because the prepositional phrase cannot be the direct object of *kissed.*

The filled-gap paradigm utilizes this predictive property of the parser by manipulating the structure following the verb, creating a violation of the parser's expectations. Specifically, we replaced the gap with an overt noun phrase, as in "The zebra that the hippo kissed the camel on the nose." When the filled-gap noun phrase "the camel" is encountered, the parser predictions about structure following the gap are violated. This expectancy violation will be indexed by an electrophysiologically "surprise" response. In particular, Hestvik et al. (2007), found that this expectancy violation generated an Early Left Anterior Negativity (eLAN) ERP in adult subjects. The eLAN is a left-anterior negative deflection in the on-going voltage measured at the scalp, and peaks as early as between 100-200 milliseconds after the onset of the violation. It has been shown to be elicited by violations of phrase structure and violations of syntactic category expectations (Friederici et al., 1996; Hahne and Friederici, 1999; Hahne and Friederici, 1999). We therefore take an eLAN response to the filled-gap noun phrase as a direct indication that the parser has attempted to fill a gap before that noun phrase.

Note that the sentences in the filled-gap paradigm of Hestvik et al (2007) could, at the temporal point of the filled gap, in principle turn out to have a grammatical continuation, as in "The zebra that the hippo kissed the camel for" or "The zebra that the hippo kissed, the camel liked." This situation would parallel that of the classical "filled gap effect": Clifton and Frazier (1989) discovered that when subjects read sentences like "What did the cautious old man whisper to his fiancé about during the movie last night?", a significant slowdown in reading time occurred during the region after the preposition *about*. The explanation is that subjects initially postulate a gap after *whisper*; then once *about* is encountered, the parser realizes that the gap was miscalculated and should have been after *about*. This leads to a slowdown caused by reanalysis and re-creating a gap representation in the correct place. In other words, subjects experience a temporary Garden Path effect (Crain and Fodor, 1985; Stowe, 1986). Thus, at the filled gap position, our stimuli sentences are theoretically not ungrammatical but similar to classical filled gap effect sentences. Some researchers have attempted to identify the eLAN with grammaticality rather than probability (Friederici, 1997; Hahne and Friederici, 1999), whereas other researchers argue that it is probabilistic in nature (Kaan and Swaab, 2003).

However, it is orthogonal to the purposes of the current study whether the eLAN response reflects absolute ungrammaticality or low probability. The presence of an eLAN at the onset of the ungrammatically filled gap noun phrase serves as an indication that a subject's comprehension system did not expect to encounter that noun phrase, because a gap consistent structure was expected. The expectancy violation ERP can therefore in turn be used to test our hypotheses about SLI children's gap-filling. If SLI children do not compute filler-gap representations at all, as entailed by the knowledge-deficit model, no eLAN is expected at the ungrammatically filled gap. On the other hand, if SLI children do attempt to fill the gap in conformance to a normal grammar (as expected under a processing-deficit model), they should exhibit the same eLAN as typically developing children. Alternatively, if they compute expectations about gap consistent structure but do so with a delayed latency, a delayed ERP violation response to the filled gap is expected. Therefore, the filled gap ERP paradigm will allow us to distinguish between the competence deficit and the performance deficit models.

In a recent study, Fonteneau and Lely (2008) conducted an ERP study of SLI children's processing of Wh-questions They compared (4a) to (4b) with ERPs timelocked to the boldfaced noun:

(4) (a) Who did Barbe push the **clown** into the wall?

(b) Who did Barbe push the **ball** into?

The paradigm was based on the authors' observations that the SLI children in their cohort typically produced Wh-questions which involved repeating the questioned noun in the gap-position. The authors report that normal controls exhibited an ELAN when the two boldfaced nouns where compared, whereas the SLI group exhibited an N400-like response. However, this study suffers from multiple problems: First, notice that "push" subcategorizes for a NP, so that "the clown" is actually fully grammatical in this position. Rather, the ungrammaticality is caused by the last NP, "the wall". Therefore, even though an ELAN was observed in the two control groups (and in the adults), it is not clear whether this can be interpreted as a violation of syntactic expectations caused by dependency completion. Secondly, the group of SLI children in the study was composed of subjects ranging from 10 years old to 21 years old. Given that ERP components typically change with age, it is highly uncertain what group their results could generalize to. Nevertheless, the study potentially point to the same conclusions as the current study's result, which focused on processing of relative clauses rather than Wh-questions, to which we now turn.

2. Method

2.1 Subjects

Thirty-one children were enrolled in the study. One child was subsequently excluded because he was diagnosed with ADHD and did not complete language screening tests. Of the remaining 30 children,13 children (9 boys, 4 girls) were classified as having Specific Language Impairment (SLI), mean age 10;3 [years: months] (range 9;1 to 12;5), and 17 children had typical language development (TLD) (10 boys, 7 girls) with a mean age of 10;5 (range 8;5-12;3). One boy and one girl in each group were left-handed. All children had English as their first language. All SLI children were receiving speech pathology services in school at the time of the study. As determined by parent questionnaires, none of the SLI children had any history of frank neurological impairments, psychological or emotional disorders, attention deficit disorders or other neuro-developmental disorders. None of the children had phonological or articulatory deficits, as determined by laboratory speech pathologists.

The children were tested on a battery of screening tests, including *Clinical Evaluation of Language Fundamentals, Fourth Edition* (Semel et al., 2004), Test of Non-Verbal Intelligence, Third Edition (Brown et al., 1997), and the Peabody Picture Vocabulary Test, Fourth Edition (Dunn and Dunn, 2007), and pure-tone hearing screening at 20dB. All children had normal hearing. Each SLI child scored at least 1.25 *SD* below the mean on at least two of the four core subtests of CELF-4. The mean expressive score on CELF was below 1.5 *SD* of the mean for the SLI group. The TD children all scored within 1 *SD* from the mean on CELF-4 (cf. Table 1). Both SLI and TD children scored within normal limits on the TONI-3. The TD children scored within normal limits on the PPVT-4; SLI children scored somewhat lower, but their mean was still within 1 SD of the mean. Table 1 summarizes the screening test scores of the two groups.

(INSERT TABLE 1 HERE)

A one-way ANOVA with 4 levels of screening tests (CELF-R, CELF-E, PPVT, TONI) was conducted to confirm that the groups differed and were equated along the relevant criteria for SLI.As expected, a main effect of group ($F(1,25)=21.1$, $p<.0005$), screening test $F(3,75)=6.6$, $p<.0005$), and a significant group x test interaction ($F(3,75)=6.2$, $p<.001$) were observed. Post-hoc Scheffé test showed that the two groups did not differ in IQ ($p=0.96$) or vocabulary score ($p=.83$), but differed significantly on the language screening tests CELF-Expressive ($p<.0005$) and CELF-Receptive ($p <.05$).

2.2 Experimental design and stimuli

The ungrammatically filled gap sentences were constructed by taking a grammatical relative clause in a sentence like *The zebra that the hippo kissed in the nose* (cf. (a) in Table 2) and inserting a distinct noun phrase in the predicted gap position, as in *The zebra that the hippo kissed <u>the camel</u> on the nose ran far away* (b) (UNGRAM condition). In addition, in order to derive an ERP, the EEG during processing of the unexpected noun phrase must be compared to EEG during processing of the same noun phrase in a similar context but where it is grammatical. (Hestvik et al., 2007), the control condition was constructed by embedding the clausal part of the relative clause as an embedded clause, as in *The lion said <u>that the hippo kissed the camel on the nose</u> and then ran far away*, cf. Table 2, row (d). However, the comparison between

UNGRAMM and OBJECT is confounded by the presence of a long-distance dependency in the former but no such dependency in the latter. The long-distance dependency itself is known to incur so-called "long negativity" during the period that the filler is kept in working memory until a gap is found (Fiebach et al., 2001; Fiebach et al., 2002), and this slow potential could complicate comparison at the critical noun phrase. In Hestvik et al. (2007) this was shown by analysis to not have an effect, but in order to completely remove this confound, we added a new control condition (ADJUNCT) by using a relativized time adverb instead of a relativized object, as in *The weekend that the hippo kissed the camel on the nose ran far away,* cf. Table 2, row (c). We follow standard syntactic theory and assume that this adverb leave a gap that is adjacent to the verb but in a VP-peripheral adjunct position. This will be taken as the central comparison in the current study, as the only difference between the critical noun phrase in UNGRAM vs. ADJUNCT is whether it results in ungrammaticality or not. The OBJECT and TRACE conditions were used as fillers in the current study, because both stimuli reduce the subjects ability to predict whether a sentence will be in the ungrammatical condition.

(INSERT TABLE 2 HERE)

If children with SLI do not attempt to construct a gap after e.g. *kissed* in Table 2, then their brain response to *the camel* in the UNGRAM condition should not be different from the brain response to the same *the camel* in the ADJUNCT condition. This logic can be illustrated with reference to Figure 1, where it can be seen that both conditions contain a dependency relation spanning the verb and the position immediately after the verb.

(INSERT FIGURE 1 HERE)

The only difference is that a gap should be inserted before *the camel* in UNGRAMM, which should result in a differential brain response to *the camel* in the two conditions: Unexpected in UNGRAMM, and expected in ADJUNCT, because a relativized time adverbial is not predicted to be related to a gap in the direct object position. (It is critical for this logic that “the weekend” indeed is immediately determined to function as an adverb rather than an object, based on its lexical semantics and the verb’s selectional restrictions).Finally, the TRACE condition was included as a filler condition to reduce the expectancy that a relative clause would always contain a filled gap.

Thus, the current stimuli were identical to the stimuli in Hestvik et al. (2007), except for the ADJUNCT condition, which was added in the current experiment. The full stimulus set (in turn derived from the stimulus design of Love (2007)) was constructed as follows: First, 32 sentences of the form in Table 2 were constructed in the TRACE condition. For each sentence in the TRACE condition, a corresponding sentence in the UNGRAMM, ADJUNCT and OBJECT condition was constructed. The same argument noun phrases and verbs were used in each set of four sentences. A second list of stimuli was then constructed by switching the agent and patient noun phrase in each verb with that of another verb, and by making other changes to the post-verbal continuation part of the sentence. The two lists thus comprised 64 unique sentences in each condition. An additional list of 64 filler sentences was constructed from 38 additional sentences that did not bear a syntactic resemblance to the relative clause sentences (thus presenting most twice). The total set of stimuli thus consisted of 320 sentences, divided into two major lists of 5 lists each.

Each set of four sentences was furthermore matched to a single comprehension question, asking about the stimulus sentence just heard. The 64 comprehension questions where of four types: Object Wh-questions ("Who did the alligator tap?"), subject Wh-questions ("Who bumped the duck?"), Yes-No questions ("Did the hippo kiss the camel?") and a set of "easy" non-content Yes-No questions ("Did you hear the word 'road'?"). Question type was counterbalanced with experimental condition type of the stimulus sentences. Because each stimulus sentence was followed by a question, every question was asked four times over the entire experiment. In addition, the filler sentences were followed by exclamations like "Is that so?" Each question was matched with two picture response options. One picture represented an object or character. The other picture represented a question mark. Subjects were instructed to select the depicted object if it represented the answer, or the question mark if not. Half the trials presented a picture depicting the correct answer, the other half required the choice of the missing answer. All the UNGRAMM sentences were matched to the missing answer option (so as not to ask a comprehension question about an ungrammatical sentence). The stimulus sentences and questions were digitally recorded using 16bit resolution and 22050kHz sampling rate, with the sentences spoken by one female speaker, and the questions by a different female speaker.

2.3 Experimental procedure

Participants were fitted with a 65-channel Geodesic Sensor Net (Electrical Geodesics) with silver/silver-chloride (Ag/AgCL) plated electrodes encased in electrolyte-wetted sponges. One electrode was placed under each eye to monitor eye movements and eye blinks. Participants were then comfortably seated in a sound- and electrically-shielded booth that was dimly lit.

They faced a computer screen positioned at eye level at a distance of 70 cm. The stimulus presentation and behavioral response collection was controlled by a PC with E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) and a Serial Response Box from Psychology Software Tools. The sentences and questions were presented at 65 dB SPL with free-field loudspeakers placed behind and in front of the subject. Participants were instructed to position the index and fourth finger of their right hand on a response box with labeled buttons.

Each experimental trial proceeded as follows: First, a picture of an eye, serving as a fixation aid and a reminder not to blink, appeared in the center of the computer screen for 1000 ms. This was followed by presentation of the stimulus sentence, with the fixation eye picture remaining on the screen during the presentation. Next, a 1000 ms gap of silence ensued, followed by auditory presentation of a question. The two response options were then depicted for 7000 ms. One button represented each depicted response option. Following a response, or the 7000 ms allowed, a 1000 ms pause ensued prior to the next trial. Accuracy feedback was provided after each question, as well as the cumulative accuracy. Subjects were encouraged to maintain high cumulative accuracy, which were intended to focus their attention on the task at hand.

The experiment started with a set of practice trials. Each subject then heard all the stimuli in two consecutive sessions. Half the subjects heard the items in List 1 first, the other half heard the items in List 2 first. Each session was further divided into four blocks of 32 trials, randomly drawn from each of the four condition sub-lists in a given session list, as well as 32 filler items. Alternation of filler versus test item was also randomized. Short breaks were given between each block, and a longer break between the two list sessions. Subjects were not told in advance that some sentences would be ungrammatical, but simply to listen for content in order to answer comprehension questions. Upon completion of the experiment, each subject was rewarded with their choice of a small toy or sticker. Parents/guardians were reimbursed for their time at $15/hour including limited travel/parking expenses. The entire recording session took between 1 ½ to 2 hours.

2.4 EEG recording and data processing

EEG was continuously recorded with a sampling rate of 200 Hz, referenced to Cz. Electrode impedances were kept below 60 kΩ which is acceptable for high impedance amplifiers (Ferree, Luu, Russell, & Tucker, 2001). EEG was amplified using a hardware band-pass filter (0.1 – 41.3Hz) and digitized using 12 bits resolution. After recording, the continuous raw EEG was segmented into 1000 ms epochs, using a 200 ms pre-stimulus baseline and a 1000 ms epoch duration. ERPs were time-locked to the onset of the first word following the relative clause verb. The 256 experimental trials per subject were then submitted to automatic artifact detection procedures. Using Netstation software, a channel in a single recording was marked as a bad channel if the fast average amplitude exceeded 200 μV; if the differential amplitude exceeded 100 μV; or if it had zero variance. A channel was considered to be a bad channel in all trials if it was a bad channel in 20 percent of trials. A trial was excluded if it contained more than 10 bad channels, or if it contained lateral eye movements resulting in amplitudes greater than ±70 μV Bad channels were deleted and replaced with data using the spherical spline interpolation, as long as they were surrounded by channels with good data. Each trial was then baseline corrected by subtracting the mean voltage during the 200ms period before the time-lock. Using the remaining trials, an average voltage per condition was computed, and re-referenced to the average voltage. Note that we removed all trials with eye-blink activity. Various alternative methods for subtracting eye-blink activity from single trials exist, such as Independent Component Analysis followed by subtraction of eye-blink components (Bell and Sejnowski, 1995). However, we have found that ICA eye-blink subtraction also attenuates the experimental effect size of the LAN and especially the ELAN component, which are located at anterior inferior electrode sites. We therefore chose not to use eye-blink subtraction and instead removed trials with eye-blinks. In addition to the increased artifacts typically observed with children, this resulted in a relatively high number of trials lost to artifacts.

3. Results

3.1 Behavioral data

The mean proportion correct answers (standard deviations in parenthesis) on comprehension questions by condition is given in Table 3:

(INSERT TABLE 3 HERE)

These differences were statistically analyzed with a mixed factorial repeated measures ANOVA with the between-subject factor GROUP (TD vs. SLI) and within-subject factor CONDITION (four levels) with percent correctly answered questions as the dependent measure. This revealed a main effect of group ($F(1,28)=9.1$, $p < .01$), and a main effect of condition ($F(3,84) = 13.3$, $p < .001$). However, there was no interaction between group and condition. In other words, the difference in accuracy between SLI and TD was statistically the same for all experimental conditions; SLI children's overall lower accuracy was not driven by significantly poorer comprehension of sentences with long distance dependencies (e.g., ADJUNCT and TRACE vs. OBJECT).

Accuracy of comprehension questions by question type was also analyzed in this way, to assess whether SLI children were poorer at comprehending object questions (which itself contains a filler-gap dependency) than subject questions and Yes/No-questions:

(INSERT TABLE 4 HERE)

A mixed factorial repeated measures ANOVA with the between-subject factor GROUP (TD. vs SLI) and within-subject factor QTYPE (question type) with percent correctly answered questions as the dependent measure revealed a main effect of group ($F(1,28)=9.3$, $p < .01$), a main effect of Condition ($F(2,56) = 28.1$, $p < .001$) such that object questions had lower accuracy than subject questions, which in turn had lower accuracy than Yes/No-question. However, there was no interaction between the two factors. In other words, although SLI children had overall lower accuracy on comprehension questions, the two groups patterned the same way with respect to question type.

3.2 ERP results

After artifact detection and removal as described in section 2.4, the two groups did not differ in terms on number of trials lots to artifacts. The mean proportion of good trials out of a total of 256 trials for the TD group was 55% (SD = 19%, range: 18%-84%), compared to 56% (SD = 18%, range: 35%-99%) for the SLI group. The high number of artifacts is to be expected with pediatric subject populations. Due to the clinical nature of the target group and the difficulty of recruiting subjects, we relaxed trial count requirements and included all subject in the data analysis. Difference wave forms where constructed by subtracting the voltage values in the ADJUNCT condition from the voltage values in the UNGRAMM condition. Figure 2, left panel shows topographical plots of the difference wave over time for the TD control group in 100ms increments from baseline, and clearly shows an anterior negativity developing from 100ms to about 500ms. Figure 2, right panel shows the comparable time series for the SLI group. As is evident, the SLI children showed no anterior negativity to the filled gap in the same early temporal and spatial region.

(INSERT FIGURE 2 HERE)

In order to determine statistical significance of the anterior negativity effect, the set of anterior inferior electrodes were averaged together. Electrodes can be grouped together on the basis of ANTERIORITY (anterior vs. posterior electrodes), LATERALITY (left vs. right hemisphere, excluding the midline electrodes), and DORSALITY (inferior vs. superior electrodes) (Dien & Santuzzi, 2005). Figure 3 shows the eight resulting electrode regions for the 64 electrode montage used in the recording (see Luu & Ferree (2000) for the correspondence between electrode placements and the International 10-10 system).

(INSERT FIGURE 3 HERE)

Given that visual inspection of the topographical plots revealed that the effect occurred bilaterally in the anterior inferior electrodes, we assessed significance of the effect by averaging, for each subject, the mean voltage values of the left and right anterior inferior regions, including the two midline electrodes 7 and 10; furthermore, we included the eye electrodes 64 and 63 from the artifact decontaminated data, as cortical activity is registered in these electrodes as well. Mean voltages were then calculated for all the subjects in 100ms time bins starting from a 100ms baseline (i.e., -100 to 0ms), and ten 100ms time bins for the rest of the epoch. The baseline period was included in order to verify that no condition effects were present during the baseline period. These dependent measures were then submitted to a mixed factorial repeated measures ANOVA, with group as the between-subject factor, and TIME (11 levels) and CONDITION (adjunct vs. ungramm) as the within-subject factor. This resulted in a CONDITION x GROUP interaction ($F(1,28)=4.76$, $p =.037$), and a marginally significant TIME x CONDITION x GROUP interaction ($F(10,280)=1.83$, $p = 0.054$). The CONDITION x GROUP interaction was caused by the UNGRAMM condition being more negative that the ADJUNCT condition for the TD group, whereas the opposite was the case for the SLI group. Given that we had a priori predictions about group differences in the filled gap effect, the CONDITION x GROUP interaction was followed up with contrast analysis of the difference between UNGRAMM and ADJUNCT for each group. For the TD children, the difference was marginally significant ($t=1.76$, $p = 0.088$), however, given that we had a prediction that the difference should go in one direction with UNGRAMM being more negative than ADJUNCT, the statistic can be interpreted with a one-tailed probability of $p=.044$. Furthermore, when the ADJUNCT vs. UNGRAMM contrast was tested separately in each time window for the TD group, it was significant in the 0-100ms time window ($t=2.1$, two-tailed $p=0.044$), as well as in the 100-200ms time window ($t=2.2$, $p=.03$), but not in the 200-300ms time window ($t=1.6$, $p=0.11$) nor the 300-400ms window ($t=1.56$, $p=.12$). However, in the 500-600ms time window, the contrast analysis again revealed a significant effect (t=2.1, p=.04). Subsequent time windows showed no significant differences. Thus, the statistical analysis confirmed that the anterior negativity was significant during the first 200ms after onset of the filled gap noun phrase in the TD children (consistent with the temporal time course of the eLAN), as well as during 500-600ms (consistent with the temporal time course of a LAN). Figure 4 illustrates the bilateral nature of the anterior negativity in the TD group:

(INSERT FIGURE 4 HERE)

Turning to the SLI group, the UNGRAMM vs. ADJUNCT contrast did not significantly differ. In other words, the SLI children exhibited no ERP component related to the ungrammaticality of the filled gap noun phrase (as is evident upon inspection of the difference waveform topoplot in Figure 2.).

However, inspection of the difference waveform plots also show a late right-posterior negativity to the filled gap noun phrase in the SLI children, cf. the 500ms, 600ms and 700ms time points in Figure 2. Using the electrode regions defined in Figure 3, we examined whether this effect was significant in the right posterior inferior electrode group, by calculating mean voltages in each 100ms time bins from 400-1000ms. This revealed a CONDITION x GROUP interaction such that UNGRAMM condition was more negative than the ADJUNCT condition in the SLI group, but no such difference in the TD group. The interaction was followed up with Fisher Least Significant Difference post hoc test revealed that the contrast was marginally significant for the SLI group (pooled MSE=7.6, df-51.56, p=0.07).

However, note that the *a priori* defined symmetrical electrode regions do not correspond clearly to the actual spatial distribution of the ERP. In order to better assess the significance of this unpredicted effect, we therefore conducted a spatial Principal Component Analysis (PCA). PCA decomposes the overall variance into underlying variance components based on the strength of covariance pattern among electrodes. Following the procedures developed by J. Dien (Dien, 1998; Dien & Frishkoff, 2005; Dien,Frishkoff, Cerbone, & Tucker, 2003; Dien, Spencer, & Donchin, 2003, 2004, 2005; Spencer, Dien, & Donchin, 1999,2001), a spatial Principal Component Analysis (PCA) analysis was conducted. The input was a data matrix consisting of 65 columns, one for each channel. The rows were 12 subjects x 4 conditions x 300 time points for each of the 65 channels (one SLI subject's data could not be included due to a recording error which created a data format incompatibility for the analysis software for this subject). The covariance among channel was computed, and the resulting relationship matrix was decomposed using PCA (eigenvalue decomposition). To determine how many factors to retain, we used a combination of Rule N (a version of the parallel test, Dien (1998)), Rule A4 (Preisendorfer & Mobley, 1988), and North's rule (North, Bell, Cahalan, & Moeng, 1982), which resulted in 11 factors to retain. The factors were then rotated using the covariance matrix (without Kaiser normalization) to simple structure, using PROMAX (*k*=3) (Hendrickson & White, 1964; Richman, 1986; Tataryn, Wood, & Gorsuch, 1999). After rotation, the 11 factors accounted for 89.5% of the original variance. Visual inspection of the factor loadings back-projected into electrode space and scaled in microvolts, showed that one spatial factor had a consistent pattern of activation across right posterior electrodes, corresponding to the distribution of the late posterior effect in the raw voltage data, and that the UNGRAMM condition was more negative going than all the other conditions. This corresponds to the observed right-posterior negativity in the difference waveforms in Fig. 2.

The spatial factor scores were then averaged for each subject and condition over 100ms time bins between 400 and 900ms (this time window was based on visual inspection of the time course of the ERP), and the resulting time-averaged factor scores were used as dependent measures in a repeated measures ANOVA (Dien & Frishkoff, 2005; Spencer et al., 2001). This is comparable to using averaged time windows of raw voltage data electrode regions, except every electrode is used, and each electrode is weighted by its contribution to the spatial PCA factor. The repeated measures ANOVA with CONDITION (four levels) and TIME (five 100ms time bins between 400 and 900ms) resulted in a significant effect of condition ($F(3,33)=3.14$, $p=.038$, Greenhouse-Geiser $\varepsilon=0.79$, $p=0.05$). Because we had general a priori predictions about the specific contrast between ADJUNCT and UNGRAM, an orthogonal contrast analysis comparing the two critical condition (averaging over the entire 400-900ms time windows) found a significant difference ($t=2.84$, $p = .016$). The general difference is illustrated in Figure 5, showing the peak electrode from spatial PCA back-projected into voltage space.

(INSERT FIGURE 5 HERE)

4. Discussion

Given the previous literature, it was expected that SLI children would have poorer performance on comprehension questions following the ADJUNCT and TRACE condition, than the OBJECT condition, as the former two contain long-distance dependency. As for the UNGRAMM condition, we chose not to interpret the behavioral performance on this condition, as the sentences are ungrammatical, and meaningful comprehension questions cannot be asked (which is why this condition always had the “no answer” option as the correct response). Furthermore, group differences in accuracy was expected to depend on whether the question was an object Wh-question, which itself involves a filler-gap dependency, vs. subject Wh-questions and Yes/No-questions. However, analysis of the behavioral off-line data related to comprehension questions revealed only a main effect of group, such that SLI children overall had poorer accuracy. There were also main effects of experimental condition on comprehension accuracy, such that object questions had lower accuracy than subject questions and Yes/No-questions, but again, there was no interaction with group. The behavioral data therefore only shows that SLI children show overall poorer accuracy, but also that this depression in comprehension is not related to filler-gap dependencies. This inference arises from the lack of interaction between group and condition, and group and question type, on the accuracy outcome. In other words, the different accuracy patterns depending on whether condition or comprehension question contained a filler-gap dependency, was the same for both groups of children. There was therefore no accuracy depression specifically related to filler-gap dependencies in the off-line, behavioral data. This matches the finding in our previous behavioral, cross-modal priming study of antecedent reactivation (Hestvik et al., 2010), where we observed that whereas SLI children failed to show antecedent reactivation immediately following the verb, they did in fact not differ significantly from the TD control group children on comprehension accuracy, which suggested that they were computing the correct representation and hence correct meaning of the stimuli sentences. However, lack of antecedent reactivation immediately following the verb therefore suggested delayed filler-gap computation.

Turning to the ERP data, it is therefore even more interesting to observe a clear group difference in the filled gap ERP response. The typical children exhibited an anterior negativity that developed and peaked within the first 200ms onset of the ungrammatical noun phrase. Although the negativity was bilateral rather than left lateralized in the average voltage data, this effect is consistent with, and largely similar to the adult filled-gap ERP response in the same paradigm, reported in Hestvik et al. (2007). In the SLI group, there was no statistically significant condition effect in the same spatiotemporal region. Thus, there was no evidence that the SLI children, despite showing fairly good comprehension of the stimuli sentences, immediately predicted a gap after the verb with the same latency as the typical children. This also converges with the previous findings of lack of filler reactivation immediately following the verb: If the SLI children have not predicted the occurrence of gap, and filled it immediately after encountering the verb, they would therefore also not be “surprised” to encounter a noun phrase occupying the same position. Thus, the fact that no immediate anterior negativity response is generated is consistent with a lack of immediate filler reactivation at the gap. In other words, the ERP results during the 200ms immediately following the verb converge with the previous behavioral findings using cross-modal lexical priming.

What are the neurobiological implications of the current finding? The current findings can be tied to the neural processing theory of Friederici (2002). According to Friederici, the brain regions involved in syntactic phrase structure building processes and which are generating the eLAN response is the anterior portion of the superior temporal gyrus and the frontal operculum (Friederici et al., 2000; Friederici et al., 2006). The eLAN, generated by this region, reflects highly automatic processes of initial structure building (as opposed to later, controlled processes indexed by P600). In normal children, the ELAN is present from age 7 according to (Hahne et al., 2004). We can infer that these brain regions in the SLI children do not respond to the filled gap violations in the current experiment. The absence of early anterior negativity suggests that this brain region is not involved in automatic processing related to structure building and structural predictions in SLI children. However, this is speculative, because, as Friederici (2006) points out, little is known about structural abnormalities and functional brain areas in SLI. Ullman and Pierpont (2005) suggest that SLI is related to a deficit in procedural memory, via abnormal development of a network of interconnected structures rooted in frontal/basal-ganglia circuits, but whether this could be tied to absence of early anterior negativities remains to be seen.

Could the delay or lack of predictions be related to working memory limitations in SLI children? Several studies of verbal working memory in SLI have concluded that these children score below normal in working memory span tasks (Weismer, 1996; Montgomery, 2000; Marton and Schwartz, 2003; Leonard et al., 2007). In a study of the relationship between verbal working memory span and ERP responses to filled gap in the current paradigm, Hestvik et al. (2012) found that low verbal span normal adults were delayed with about 200ms in both anterior negativity and P600 responses to filled gap, using exactly the same paradigm as reported here. However, if SLI children processed relative clauses like low verbal memory span normal adults, we should have observed merely a delayed anterior negativity, not a complete absence. The current ERP findings therefore do not point to verbal memory limitations as the explanation for the missing ERP response.

The ERP data did reveal one additional finding. The spatial PCA decomposition analysis of the SLI data did reveal a late, posterior spatial factor containing a significant difference between the UNGRAMM and ADJUNCT conditions. This ERP corresponded to a right parietal negative going waveform for the ungrammatical condition during the 400-900ms time window. Statistical analysis confirmed that this effect was significant, hence, one can infer that the SLI children's brain response show that the SLI parser knows that there is a difference between the two conditions at the filled gap noun phrase, and that this difference means that they detect the ungrammaticality of the filled gap. Further questions are raised by this finding. In particular, what is the mechanism by which this grammaticality inference arises in SLI children? In a recent paper, (Marinis and Van der Lely, 2007) suggest that SLI children may interpret filler-gap dependencies via "Direct Association" (Pickering and Barry, 1991) of the filler with the verb's argument structure, based on their observation that SLI children exhibit an early reactivation effect at the verb itself, but not at the gap position. Direct Association amounts to computing filler-gap dependencies via semantic association (or argument structure operations) rather than via a syntactic route. If this is the mechanism in SLI children, then the filled-gap NP should be interpreted as an argument structure violation, which in turn should elicit an N400 response in the model of Friederici (2002). Although the observed negativity in the current study does not have the typical distribution of an N400, this is a possibility that should be pursued in future studies of SLI gap-filling.

In conclusion, given the presence of a delayed, grammaticality-modulated ERP in the SLI children, the current findings support a processing-deficit model of SLI in the domain of filler-gap dependency computation. SLI children's processing system does distinguish between grammatical and ungrammatical sentences at the point of the gap, but do so at a delayed latency.

Acknowledgements

This research was funded by NIH/NIDCD grant 5R03DC006175 (Arild Hestvik, P.I.) and NIH/NIDCD grant 5T32DC00039-09 (R. Schwartz PI), which we gratefully acknowledge.

Figure captions

Figure 1. Comparison of the UNGRAMM and the ADJUNCT condition. The critical portions of the sentences are identical, save for the grammatical function status of the relativized noun. A long-distance dependency spans the relative clause and verb in both cases.

Figure 2: Difference wave form topo plots constructed by subtracting the ADJUNCT condition voltage values from the UNGRAMM condition voltage values; shown at 100ms intervals from baseline. Left panel: Typically developing children; right panel: Children with SLI.

Figure 3: Symmetrical scalp regions for averaging contiguous sets of electrodes.

Figure 4: The central left (EGI 14/AF7), central (EGI 10) and right (EGI 1/AF8) anterior inferior electrodes, grand average for TD children (N=17).

Figure 5. PCA factor waveform plot of the peak electrode (right posterior electrode 41) for the SLI group's (N=12) ungrammaticality-related spatial PCA factor, by experimental condition. ERP computed by back-projecting the factor into voltage space; waveform display filtered at 15Hz.

References

Bell, A. J. and T. J. Sejnowski (1995). "An information-maximization approach to blind separation and blind deconvolution." Neural Comput **7**(6): 1129-1159.

Brown, L., R. Sherbenou, et al. (1997). Test of Nonverbal Intelligence. Austin, TX, PRO-ED.

Clahsen, H. (1989). "The grammatical characterization of developmental dysphasia." Linguistics **27**: 897 920.

Clahsen, H. (2008). Chomskyan syntactic theory and language disorders. The Handbook of Clinical Linguistics. M. J. Ball, M. Perkins, N. Mueller and S. Howard. Oxford, Blackwell**:** 165-183.

Clifton, C., Jr. and L. Frazier (1989). Comprehending sentences with long-distance dependencies. Linguistic structure in language processing. G. Carlson and M. K. Tanenhaus. Dordrecht, Kluwer Academic Publishers**:** 273-317.

Crain, S. and J. D. Fodor (1985). How can grammars help parsers? Natural Language Parsing: Psychological, Computational and Theoretical Perspectives. D. Dowty, L. Karttunen and A. Zwicky. Cambridge, Cambridge University Press.

Deevy, P. and L. B. Leonard (2004). "The Comprehension of Wh-Questions in Children With Specific Language Impairment." Journal of Speech, Language, and Hearing Research **47**(4): 802-815.

Dunn, L. M. and D. M. Dunn (2007). Peabody Picture Vocabulary Test. Bloomington, MN, Pearson, Inc.

Fiebach, C., M. Schlesewsky, et al. (2002). "Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German WH-questions." Journal of Memory and Language **47**(2): 250.

Fiebach, C., M. Schlesewsky, et al. (2001). "Syntactic working memory and the establishment of filler-gap dependencies: insights from ERPs and fMRI." Journal of Psycholinguistic Research **30**(3): 321-338.

Fonteneau, E. and H. v. d. Lely (2008) "Electrical brain responses in language-impaired children reveal grammar-specific deficits." PLoS ONE **3** DOI: doi:10.1371/journal.pone.0001832.

Frazier, L. and C. Clifton (1995). Construal. Cambridge, MA, MIT Press.

Frazier, L. and G. B. Flores d'Arcais (1989). "Filler driven parsing: A study of gap filling in dutch." Journal of Memory and Language **28**(3): 331.

Friederici, A. D. (1997). "Neurophysiological aspects of language processing." Clin Neurosci **4**(2): 64-72.

Friederici, A. D. (2002). "Towards a neural basis of auditory sentence processing." Trends in Cognitive Sciences **6**(2): 78-84.

Friederici, A. D. (2006). "The Neural Basis of Language Development and Its Impairment." Neuron **52**(6): 941-952.

Friederici, A. D., J. Bahlmann, et al. (2006). "The brain differentiates human and non-human grammars: Functional localization and structural connectivity." Proceedings of the National Academy of Sciences of the United States of America **103**(7): 2458-2463.

Friederici, A. D., A. Hahne, et al. (1996). "Temporal structure of syntactic parsing: early and late event-related brain potential effects." Journal of Experimental Psychology: Learning, Memory, and Cognition **22**(5): 1219-1248.

Friederici, A. D., Y. Wang, et al. (2000). "Localization of early syntactic processes in frontal and temporal cortical areas: a magnetoencephalographic study." Human Brain Mapping **11**(1): 1-11.

Friedmann, N. and R. Novogrodsky (2004). "The acquisition of relative clause comprehension in Hebrew: a study of SLI and normal development." Journal of Child Language **31**: 661-681.

Friedmann, N. and R. Novogrodsky (2007). "Is the movement deficit in syntactic SLI related to traces or to thematic role transfer?" Brain and Language **101**(1): 50-63.

Friedmann, N. and R. Novogrodsky (2008). Subtypes of SLI: SySLI, PhoSLI, LeSLI, and PraSLI. Language acquisition and development A. Gavarró and M. J. Freitas. Newcastle, UK, Cambridge Scholars Press**:** 205-217.

Friedmann, N. and R. Novogrodsky (2011). "Which questions are most difficult to understand?: The comprehension of Wh questions in three subtypes of SLI." Lingua **121**(3): 367-382.

Gibson, E. and G. Hickok (1993). "Sentence processing with empty categories." Language and Cognitive Processes **8**: 147-161.

Hahne, A., K. Eckstein, et al. (2004). "Brain signatures of syntactic and semantic processes during children's language development." Journal of Cognitive Neuroscience **16**(7): 1302-1318.

Hahne, A. and A. D. Friederici (1999). "Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes." J Cogn Neurosci **11**(2): 194-205.

Hahne, A. and A. D. Friederici (1999). "Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes." Journal of Cognitive Neuroscience **11**(2): 194-205.

Hestvik, A., E. Bradley, et al. (2012). "Working Memory Effects of Gap-Predictions in Normal Adults: An Event-Related Potentials Study." Journal of Psycholinguistic Research **41**(6): 425-438.

Hestvik, A., N. Maxfield, et al. (2007). "Brain responses to filled gaps." Brain and Language **100**(3): 301-316.

Hestvik, A., R. Schwartz, et al. (2010). "Relative Clause Gap-Filling in Children with Specific Language Impairment." Journal of Psycholinguistic Research **39**(5): 443-456.

Joanisse, M. F. and M. S. Seidenberg (1998). "Specific Language Impairment: A deficit in grammar or processing?" Trends in Cognitive Sciences **2**(7): 240-247.

Joanisse, M. F. and M. S. Seidenberg (2003). "Phonology and syntax in specific language impairment: Evidence from a connectionist model." Brain and Language **86**: 40-56.

Kaan, E. and T. Y. Swaab (2003). "Electrophysiological evidence for serial sentence processing: a comparison between non-preferred and ungrammatical continuations." Brain Res Cogn Brain Res **17**(3): 621-635.

Kail, R. (1994). "A Method for Studying the Generalized Slowing Hypothesis in Children With Specific Language Impairment." J Speech Hear Res **37**(2): 418-421.

Kayne, R. (1994). The antisymmetry of syntax. Cambridge, Massachusetts, MIT Press.

Lahey, M., J. Edwards, et al. (2001). "Is Processing Speed Related to Severity of Language Impairment?" J Speech Lang Hear Res **44**(6): 1354-1361.

Leonard, L. B. (1998). Children with Specific Language Impairment. Cambridge, Massachusetts, MIT Press.

Leonard, L. B., S. E. Weismer, et al. (2007). "Speed of Processing, Working Memory, and Language Impairment in Children." Journal of Speech, Language & Hearing Research **50**(2): 408-428.

Love, T. (2007). "The Processing of Non-canonically Ordered Constituents in Long Distance Dependencies by Pre-school Children: a Real-time Investigation." Journal of Psycholinguistic Research **36**(3): 191-206.

Love, T. and D. Swinney (1997). Real-time processing of object-relative constructions by pre-school children. 10th Annual CUNY Conference on Human Sentence Processing, Los Angeles, CA.

Marinis, T. and H. Van der Lely (2007). "On-line processing of wh-questions in children with G-SLI and typically developing children." International Journal of Language & Communication Disorders **42**(5): 557-582.

Marton, K. and R. G. Schwartz (2003). "Working memory capacity and language processes in children with specific language impairment." Journal of Speech, Language and Hearing Research **46**(5): 1138-1153.

Miller, C. A., L. B. Leonard, et al. (2006). "Response Time in 14-Year-Olds With Language Impairment." J Speech Lang Hear Res **49**(4): 712-728.

Montgomery, J. W. (1995). "Sentence comprehension in children with specific language impairment: the role of phonological working memory." Journal of Speech, Language, and Hearing Research **38**(1): 187-199.

Montgomery, J. W. (2000). "Verbal Working Memory and Sentence Comprehension in Children With Specific Language Impairment." J Speech Lang Hear Res **43**(2): 293-308.

Montgomery, J. W. (2006). "Real-time language processing in school-age children with specific language impairment." International Journal of Language & Communication Disorders **41**(3): 275-291.

Novogrodsky, R. and N. Friedmann (2006). "The production of relative clauses in syntactic SLI: a window to the nature of the impairment." Advances in Speech-Language Pathology **8**: 364-375.

Pickering, M. J. and G. D. Barry (1991). "Sentence processing without empty categories." Language and Cognitive Processes **6**: 229-259.

Roberts, L., T. Marinis, et al. (2007). "Antecedent Priming at Trace Positions in Children's Sentence Processing." Journal of Psycholinguistic Research **36**(2): 175-188.

Schuele, C. M. and L. Nicholls (2000). "Relative clauses: evidence of continued linguistic vulnerability in children with Specific Language Impairment." Clinical Linguistics & Phonetics **14**(8): 563-585.

Schuele, C. M. and L. Tolbert (2001). "Omissions of obligatory relative markers in children with specific language impairment." Clinical Linguistics & Phonetics **15**(4): 257-274.

Semel, E., E. H. Wiig, et al. (2004). Clinical evaluation of language fundamentals. Toronto, Canada, The Psychological Corporation/A Harcourt Assessment Company.

Stavrakaki, S. (2001). "Comprehension of Reversible Relative Clauses in Specifically Language Impaired and Normally Developing Greek Children." Brain and Language **77**(3): 419-431.

Stowe, L. A. (1986). "Parsing WH-constructions: Evidence for on-line gap location." Language and Cognitive Processes **1**(3): 227 - 245.

Stowe, L. A., M. K. Tanenhaus, et al. (1991). "Filling gaps on-line: Use of lexical and semantic information in sentence processing." Language and Speech **34**(4): 319-340.

Ullman, M. T. and E. I. Pierpont (2005). "Specific language impairment is not specific to language: the procedural deficit hypothesis." Cortex **41**(3): 399-433.

van der Lely, H. (1998). "SLI in Children: Movement, Economy, and Deficits in the Computational-Syntactic System." Language Acquisition **7**(2-4): 161-192.

van der Lely, H. (in press). Grammatical-Specific Language Impairment: A Representational Deficit for Dependent Relationships: Evidence from English. Essays on Syntax, Morphology and Phonology in SLI. C. Jacubowicz, L. Nash and K. Wexler. Cambridge, Cambridge University Press.

Van der Lely, H. and J. Battell (2003). "Wh-movement in children with grammatical SLI: A test of the RDDR hypothesis." Language **79**(1): 153-181.

Weismer, S. E. (1996). "Capacity Limitations in Working Memory: The Impact on Lexical and Morphological Learning by Children with Language Impairment." Topics in Language Disorders **17**: 33-44.

Weismer, S. E., J. Evans, et al. (1999). "An Examination of Verbal Working Memory Capacity in Children With Specific Language Impairment." Journal of Speech, Language, and Hearing Research **42**(5): 1249-1260.

Windsor, J. and M. Hwang (1999). "Testing the generalized slowing hypothesis in specific language impairment." Journal of Speech, Language, and Hearing Research **42**(5): 1205-1218.

Zhang, Y. and J. Zhang (2008). "Brain responses to agreement violations of Chinese grammatical aspect." Neuroreport **19**(10): 1039-1043 1010.1097/WNR.1030b1013e328302f328314f.

Tables

Participants	Age	CELF-R	CELF-E	CELF-C	PPVT	TONI
SLI (N=13)						
M	10;3	86.75	76.50	78.83	90.50	101.00
SD	1;2	11.17	7.09	6.78	9.07	16.61
Control (N=17)						
M	10;5	106.63	106.25	101.56	99.44	109.94
SD	1;0	14.30	14.33	15.29	15.29	17.12

Table 1: Participant profiles with standard scores. CELF-C is a composite score based on CELF-E(expressive) and CELF-R(eceptive).

Condition	Pre-gap string	ERP time- lock point
(a) TRACE (filler)	The zebra that the hippo kissed	on the nose ran far away
(b) UNGRAMM	The zebra that the hippo kissed	the camel on the nose ran far away
(c)ADJUNCT (control)	The weekend that the hippo kissed	the camel on the nose he ran far away
(d) OBJECT(filler)	The lion said that the hippo kissed	the camel on the nose and then ran far away

Table 2: Illustration of the four sentence types in the experiment.

Condition	TLD	SLI
(a) TRACE	0.77 (.10)	0.64 (.16)
(b) UNGRAMM	0.64 (.08	0.60 (.10)
(c)ADJUNCT	0.78 (.10)	0.68 (.12)
(d) OBJECT	0.66 (.07)	0.58 (.10)

Table 3: Mean and standard deviations for comprehension question accuracy by group and condition.

Condition	TLD	SLI
Object questions	0.57 (0.07)	0.51 (0.12)
Subject questions	0.69 (0.07)	0.57 (0.09)
Yes/No questions	0.70 (0.09)	0.63 (0.12)

Table 4: Mean and standard deviations for comprehension question accuracy by group and question type.

Figure1

Click here to download high resolution image

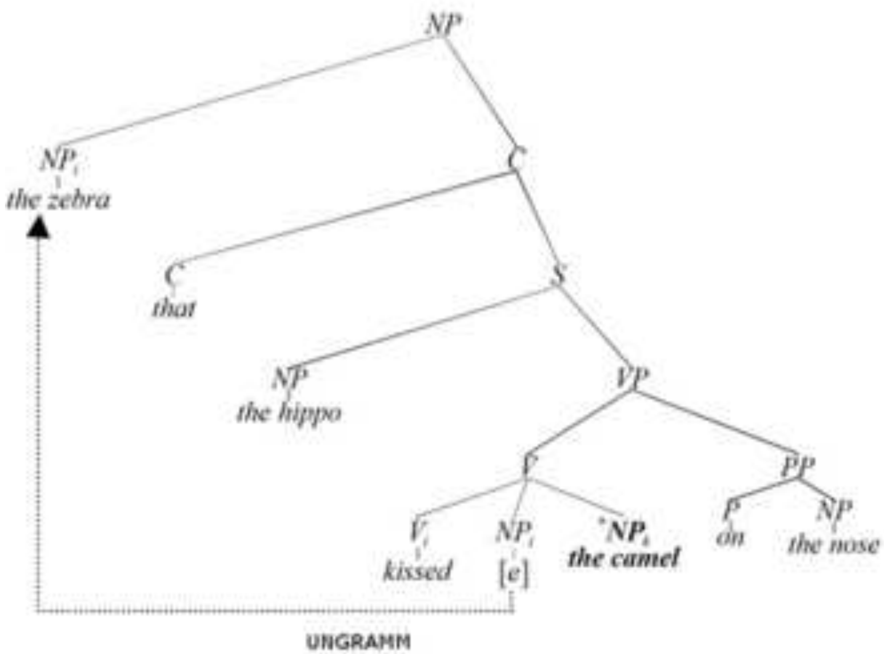

Figure2

Click here to download high resolution image

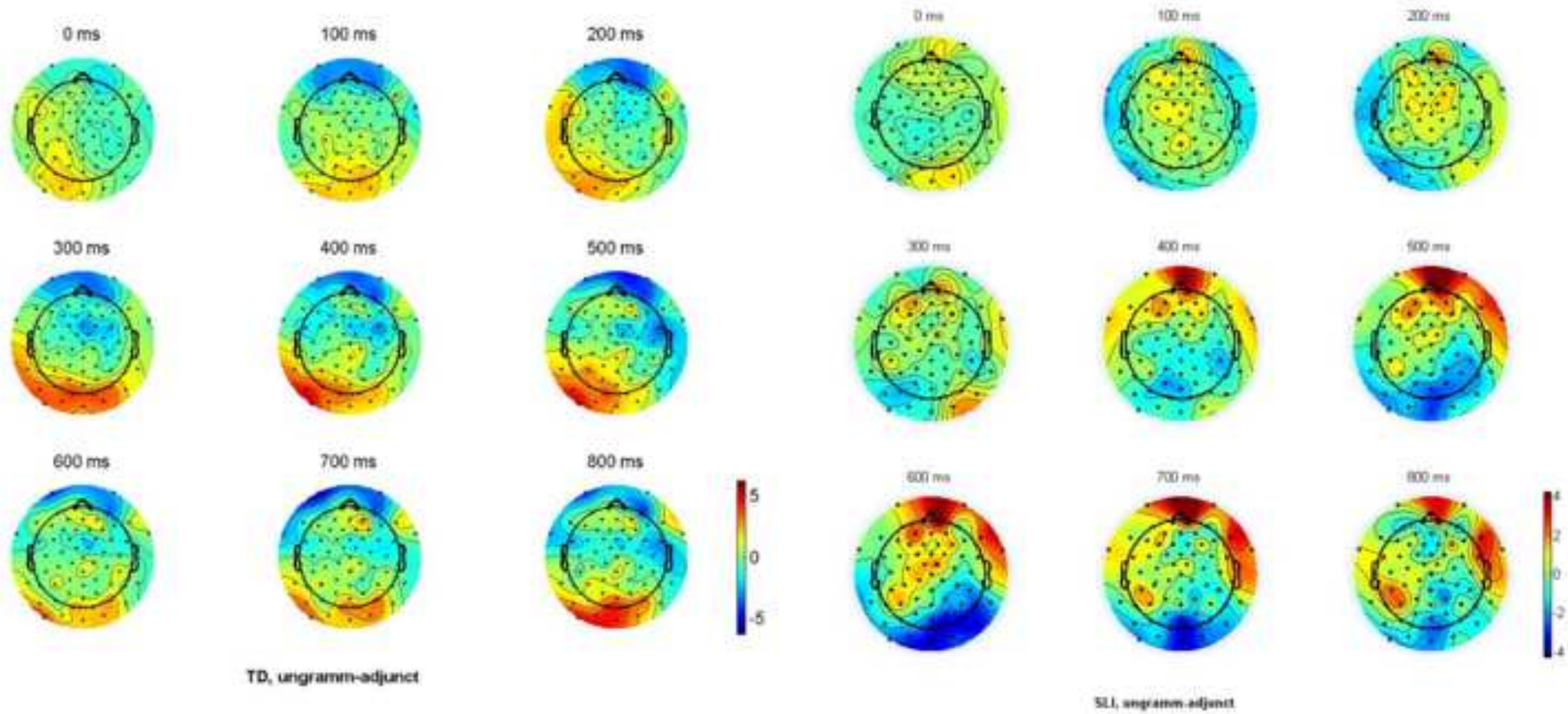

Figure3

Click here to download high resolution image

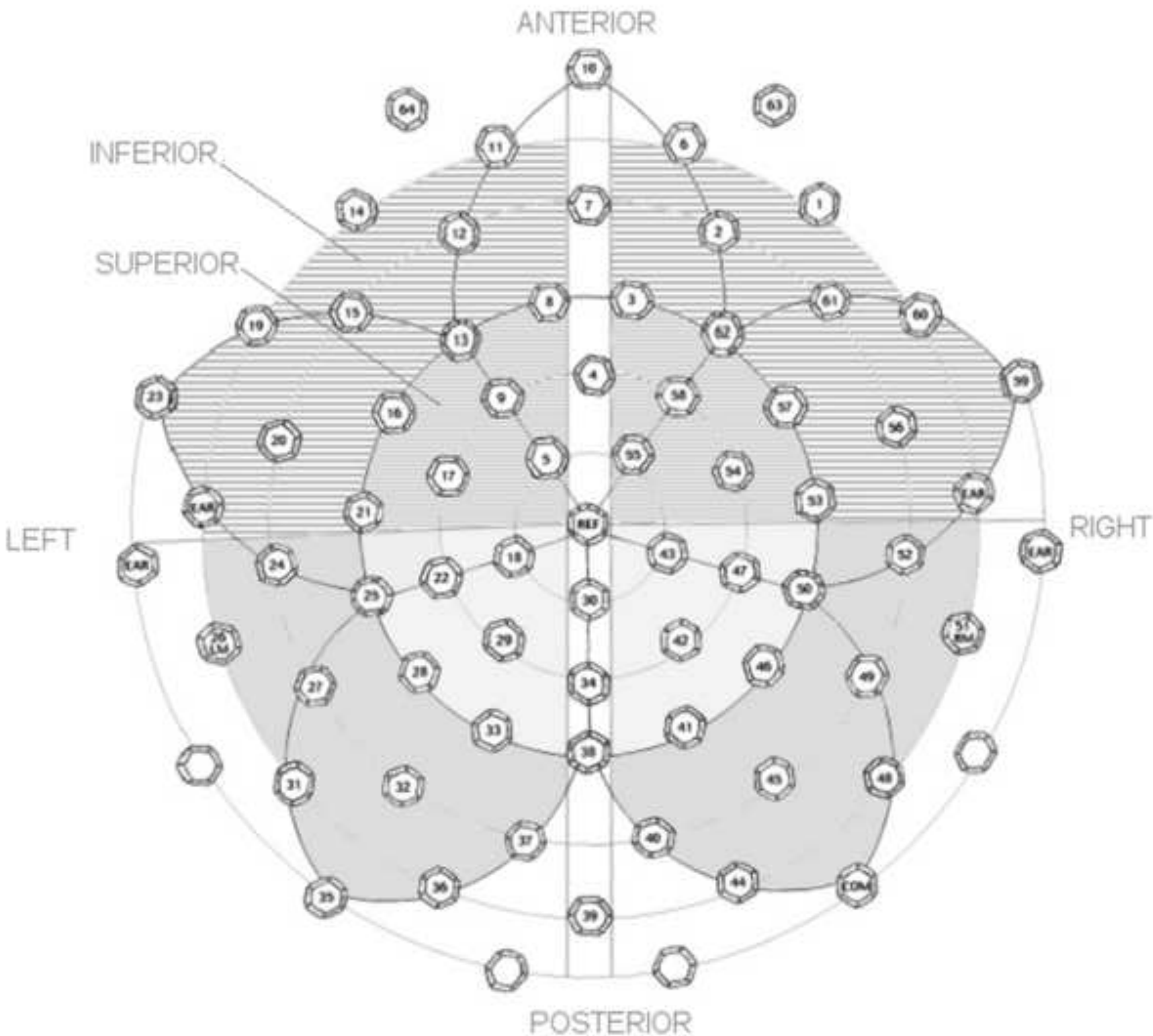

Figure4

Click here to download high resolution image

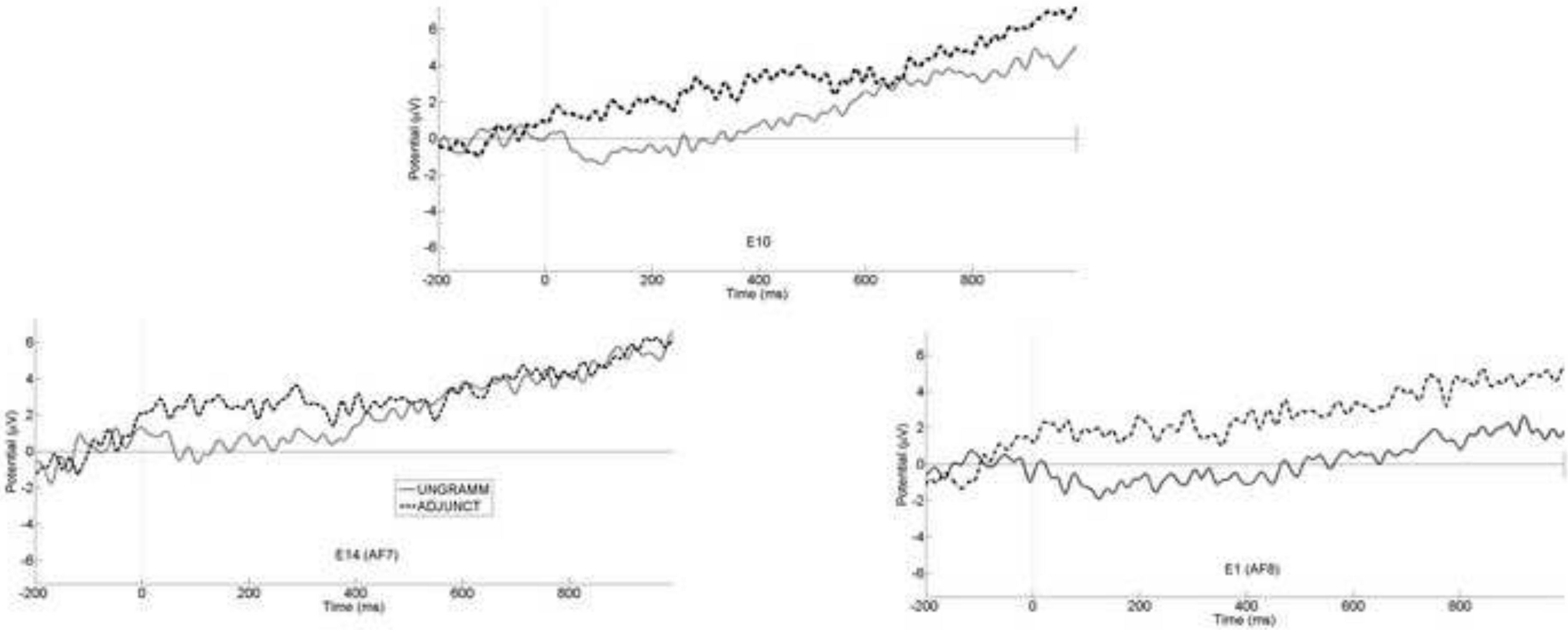

Figure5
Click here to download high resolution image