Reading & Writing (in press)

Revisiting the scrambling complexity hypothesis in sentence processing: A self-paced reading study on anomaly detection and scrambling in Hindi

Ramesh K. Mishra, Aparna Pandey, & Narayanan Srinivasan

Centre of Behavioural and Cognitive Sciences, University of Allahabad

Corresponding Author:

Dr. Narayanan Srinivasan

Centre of Behavioural and Cognitive Sciences

University of Allahabad

Allahabad 211002

India

Email: nsrini@cbcs.ac.in

Phone/Fax: +915322460738

Abstract

The scrambling complexity hypothesis based on working memory or locality accounts as well as syntactic accounts have proposed that processing a scrambled structure is difficult. However, the locus of this difficulty in sentence processing remains debatable. Several studies on multiple languages have explored the effect of scrambling on sentence processing and not all languages have shown an advantage for the canonical word order. Using a self-paced reading paradigm, we studied the effect of scrambling on semantic anomaly detection in Hindi sentence comprehension employing three word order types. Reading times on critical verbs, judgment latency, and error rates showed significant effect of word order type. The results further revealed significant interactions between word order and anomaly type. The patterns of results suggest that the canonical word order does not necessarily have a processing advantage in terms of speed and accuracy over non-canonical orders and do not provide support to sentence processing accounts that assume an advantage for canonical structures. The results indicate that processing speed depends on the distance between the subject and the verb, thus supporting a locality dependent working memory based model of sentence processing. The results provide evidence for the role of specific cognitive processes in Hindi sentence processing with further implications for language and literacy acquisition in Hindi.

Key words: Scrambling; working memory; canonicity; semantic anomaly; self-paced reading; Hindi; sentence processing

Revisiting the scrambling complexity hypothesis in sentence processing: A self-paced reading study on anomaly detection and scrambling in Hindi

Introduction

Natural languages exhibit different types of linearization in word orders. Transformational generative approaches to syntax assumes that sentences in all languages have a base form, that is a canonical word order like Subject-Object-Verb (SOV) for Hindi, and the other word orders are derived from this base structure involving movement and transformation (Chomsky, 1981). A canonical word order has been defined as the most statistically frequent (Kayne, 1994) as well as contextually neutral word order (Comrie, 1981). They could be Subject-Verb-Object (SVO) as in English or SOV as in Hindi and Japanese. The process where a speaker uses a non-canonical over canonical word order has been termed as scrambling.

Theoretical accounts explaining the effect of scrambling on sentence processing are inconclusive in spite of studies in different languages (for a review see Sekerina, 1997). The effects of scrambling on sentence processing, that is on error rates or judgment latencies have been used to argue for different theoretical models of sentence processing itself. Both syntactic (Frazier, 1995) and memory-based retrieval accounts of sentence processing (Lewis & Vasishth, 2005) have been proposed to account for the effects of scrambling found in natural language processing. A strictly syntactic account of word order processing predicts increased processing times for the non-canonical structures over canonical structures (Fodor, 1978; Gibson, 1998). The syntactic model of sentence parsing by Frazier and Fodor (1978) predicts such a result.

The scrambling complexity based view that processing non-canonical structures are difficult is supported by several theoretical positions. For example, the syntactic prediction locality theory (Gibson, 1998) assumes that processing non-canonical structures require greater memory resources. This demand for memory arises because of violation of the locality principle. Alternative accounts like the trace reactivation account (Nicol & Swinney, 1989) assume structures that are deleted from their canonical positions need to be kept in working memory till reintegration and this causes difficulty in processing non canonical structures. However, findings from sentence reading studies with non-canonical structures do not provide unanimous support for these predictions (Karimi, 2003). Considerable differences exist between findings in processing times for different word orders. Purely syntactic accounts have failed to predict accurately sentence-processing performance (e.g. reading times, judgment times, and errors) with canonical and non-canonical word orders in several languages.

Like Hindi, Japanese is a free word order language with SOV as the canonical pattern (Nakayama, 1999). Nakayama (1995) and Yamashita (1997) had found no difference in reading times between canonical and non-canonical sentence reading times in Japanese. Using a self-paced reading paradigm, Tamaoaka, Sakai, Kawahara, and Miyoaka (2003) studied the effect of phrase lengths and scrambling on Japanese sentence comprehension. They did not find any effect of word order or phrase length on reading times but only on error rates. They concluded that scrambling has no effect on speed of processing but non-canonical sentences are difficult to judge compared to canonical sentences. The authors pointed out that Japanese is a non-configurational language that does not need any movement, as linguistic theories would predict. They argued that, since the canonical and non-canonical word orders do not alter the meaning of a sentence, processing costs should not differ.

However, Kim, Koizumi, and Ikuta (2009) found that canonical (SOV) sentences are better processed than other non-canonical (Object-Subject-Verb: OSV) sentences in Japanese. Support for scrambling complexity also comes from a study on Basque, a free word order language (Erdocia, Laka, Mestres-Misse, & Rodriguez-Fornells, 2009). Erdocia et al. (2009) presented participants SOV (canonical in Basque) and OSV (non-canonical) sentences in a self paced reading task (Exp. 1) and measured reading times for whole sentences as well as individual words along with error rates. They found that non-canonical forms are difficult to process (larger processing times and higher error rates) indicating a significant effect of word order on sentence processing. They also observed that reading times for sentence-initial objects were faster compared to sentence-initial subjects. This result was different from those obtained with languages like Dutch (1997) and Russian (Sekerina, 1997; 2003). This unexpected result was explained by the fact that Basque is an ergative language and therefore sentence-initial objects resemble subjects in intransitive structures. Hence, subjects project a simpler structure and processing takes less time. This finding has relevance in interpreting results in Hindi as it is typologically an ergative language and has a rich inflectional case marking system.

ERP studies on scrambling also present a mixed picture (Rosler, Pechmann, Streb, Roder, & Hennighausen, 1998; Schlesewsky, Bornkessel, & Frisch, 2003; Wolff, Schlesewsky, Hirotani, & Bornkessel-Schlesewsky, 2008). Some studies have shown that scrambled sentences exhibit differential brain activity (Rosler et al., 1998; Schlesewsky et al., 2003). Rosler et al. (1998) explored the effect of word order variations on sentence comprehension in German using ERP and found that different word orders had distinct ERP signatures in terms of anterior or posterior distributions. The word orders varied from the canonical S-IO-DO sequence in terms of a linear order principle (LP-Principles: Pechmann, Uszkoreit, Engelkamp, & Zerbst, 1994).

For German, the rank order as predicted by the LP-principle was S-IO-DO smaller than S-DO-IO smaller than IO-S-DO. However, the comprehension time and decision time measures showed no such graded effects, though the canonical word order was the easiest to process.

Schlesewsky, Bornkessel, and Frisch (2003) found maximum negativity in the frontal areas for the scrambled structures. They accounted for their results suggesting that the negativity observed does not reflect a working memory demand but rather a violation of the local syntactic structure. Thus, they provided a purely syntactic account for word order processing in German as opposed to working memory based approaches to sentence processing. Importantly, they also suggested that displacement of argument structures in scrambled sentences also changed the information structure and that may have caused the effects observed in the ERP. However, currently there are no accounts that suggests a purely information structure based approach to understand processing of scrambled sentences.

The results with different languages suggest that word order variation may not always induce processing constraints and the effects of scrambling on sentence processing across languages are not consistent. Predicting the effect of scrambling on sentence processing has been difficult especially with tasks like anomaly judgment. Different theoretical ideas have been postulated for accounting the data from languages like Basque and Japanese that further need empirical examination with other languages.

Hindi is well suited to study the effects of word order variations on sentence processing because it is a flexible word order language like Japanese and German and has agreement features similar to German. These typological similarities between Hindi and German make the results comparable to some extent. In addition, there are few published studies on the effect of scrambling in Hindi. We used a self-paced reading paradigm involving semantic anomaly judgments. One important aspect of sentence processing is judging the semantic acceptability of a sentence. Judgment of a semantically anomalous sentence has been used as a sensitive task to explore the capabilities of the language processor, as in studying child language (de Villiers & de Villiers, 1972), in aggramatic aphasics (Kolk & Weijts, 1996), and in the ERP paradigm (Kuperberg et al., 2006). However, very few studies have explored the effect of sentential structure that is word order on judgment of semantic anomaly. The present study in Hindi explores the effect of word order change or scrambling on speed and accuracy of detecting semantic anomaly in a self-paced reading paradigm.

A scrambling complexity based account may suggest that anomaly detection may be faster or error rates may be less with canonical sentences. However, so far, no study that has investigated anomaly detection with canonical and non-canonical structures. The present study aims to investigate the effect of scrambling on sentence comprehension in Hindi using a self-based reading paradigm, with an aim to provide further cross-linguistic evidence on these phenomena and evaluate competing claims. Specifically, the research explored the effect of word order variation on speed and accuracy of semantic anomaly judgment in sentence processing with Hindi.

This task is different from what has been employed in previous studies on scrambling in other languages (for example in Erdocia et al., 2009). We made the sentences anomalous by making the subject impossible to be used with the verb, that is, violation of selectional restriction of the verb (Chomsky, 1965). Thus, both types of sentences were perfectly grammatical sentences from a syntactic point of view. It must be noted that several ERP studies have employed semantically anomalous sentences in studying sentence comprehension (Kutas & Hillyard, 1984). Semantically anomalous sentences have been used by child language researchers to study the development of sentence comprehension skills (De Villiers & De Villiers, 1972). Semantically anomalous sentences have also been used in studying processing of form and content during reading using eye movements (Braze, Shankweiler, Ni, & Palumbo, 2002; Ni, Fodor, Crain, & Shankweiler, 1998). One important reason for using anomalous sentences in Psycholinguistic studies of reading and sentence comprehension is to test the decision making abilities of the parser when there is increased processing cost. Importantly, sentences could be made semantically anomalous while keeping them grammatically acceptable.

We used three different word orders: SOV, SVO, and OSV with anomalous and non-anomalous sentences. In our study, we did not use a simple object but used an adjunct clause (we will continue to use the symbol 'O' rather than 'Adjunct Clause' in the rest of the paper). Adjunct clauses are secondary to the main clause in a sentence. Often adjunct clauses contain a non-finite verb with one or more noun phrases where as the main verb phrase contains the finite verb. For example in the Hindi sentence, baaris (rain) hone (causative) se (instrumental) maakan (house) giri (fell)/ House fell because of the rain, the adjunct clause is to the left of the main clause that contains the subject and the finite verb. In a free word order language like Hindi, the placement of the adjunct clause can vary either to the left or to right of the main clause, thus giving rise to the phenomena called scrambling (Dayal & Mahajan, 2004; Mohanan & Mohanan, 1994). Thus, movement of the adjunct clause could create processing differences in a free word order language.

Hindi sentences containing adjunct clauses are quite frequent in daily usage and one notices scrambled structures in which the adjunct clause is at different sentential locations. We essentially manipulated the location of the adjunct clause (Adjunct Sub Verb, Sub Adjunct Verb and Subject Verb Adjunct) while constructing the sentences used in our study. Complex sentences with adjunct classes were used to ensure enough working memory resources are required to process the sentences compared to simple sentences that may not be sufficiently sensitive to fluctuations in working memory. Sentences were made semantically anomalous, by making the subject incompatible with the main verb as a function of the causal event described by the adjunct clause. For the purposes of our study, we treat the adjunct clause containing a causative construction as a single constituent that was same in all the sentence types and its position changed. However, please note that elements inside the adjunct clause were not moved thus making this constraint consistent across stimuli. Previous psycholinguistic studies on scrambling have used sentences containing adjunct clauses of different types (Miyagawa, 2005). In addition, we used the complex sentences with adjunct clause rather than a simple object to investigate the role of cognitive processes like working memory in sentence comprehension.

A syntactic account of scrambling would predict that processing costs would be minimal for the canonical structure, that is SOV for both the anomalous and non-anomalous types and hence reading times would be faster for the canonical word order unlike Japanese. Alternatively, if the locality-based accounts of sentence processing were true, then processing with SOV sentences would be worse compared to both SVO and OSV types. This would be due to the larger distance between the subject and verb with SOV sentences leading to greater memory demands for later reintegration and judgment, in spite of the fact that this is the canonical type in Hindi. A purely structural account that gives prominence to the canonical sentence types would predict an effect in the opposite direction. Furthermore, we also predicted that processing subject initial structures would be easier compared to object initial structures. As for the error rates, we predicted that the sentence type would have a significant effect on error rates but did not make a specific prediction about the direction of the effect.

Methods

Participants

Twenty four native Hindi speakers (mean age = 21 years) provided informed consent and participated in the experiment. All the participants were right handed and had normal or corrected-to-normal vision. All the participants were naïve to the purpose of the experiment.

Stimuli

One hundred and eighty Hindi sentences were used in the experiment. Out of these, 90 were anomalous and 90 were non-anomalous sentences. Three word orders were used in the experiment, that is SOV, SVO, and OSV with 30 sentences were of each type. Anomalous sentences were prepared by changing the verbs so that they become semantically incongruent with the corresponding subjects in a sentence. The average length of the sentences was in the range of 5-7 words. All the sentences were similar in their structure according to the word order type.

An example sentence from each condition is given below (Appendix contains for a larger set of 10 example sentences from each of the conditions used in the study):

1. **SOV anomalous**

Imarat tuufan aane se bhagi

Building storm come because of ran

Building ran when the storm came

2. **SVO anomalous**

Ped	chillaye	kulhadi	se	katne par
Tree	cried	hammer	because of	cut

Tree cried when cut with hammer

3. **OSV anomalous**

Jamin	par	bahta	hua	paani	thaka
Earth	on	flow	because of	water	tired

Water got tired when flowed on ground

4. **SOV non-anomalous**

Murti	jameen par	girne	se	tutti
Idol	ground on	fall	because of	broke

Idol broke when fell o0n the ground

5. **SVO non-anomalous**

Ser	gurraya	siikari	ko	dekh	kar
Lion	roared	hunter	to	see	because of

Lion roared when it saw the hunter

6. **OSV non-anomalous**

suuar	ko	dekh	kar	kutta	bhoka
Pig	because	of	see	dog	barked

Dog barked when it saw the pig

Procedure

Participants were seated at a distance of 60 cm from the computer monitor and were instructed to read all sentences at their own pace for comprehension. A commercial software, Presentation

(Neurobehavioral Systems Inc, USA) was used to present stimuli and record the responses. The trial began with the first word of the sentence. Once participants read the word, they had to press the space bar for the next word and so on. After the last word of the sentence, the decision screen was presented asking the participants to make the judgment about the appropriateness (or the acceptability) of the sentences. Judgment time was measured as the time between the offset of the last word of each sentence until participants pressed a key. Participants had to judge the acceptability of the sentence at the end of the trial. After the judgment, a screen appeared where participants could take rest or move for next trial by pressing any key. Sentences were counter-balanced across participants and order of presentation of the sentences was random for all participants.

Results

The reading time for the critical word (the verb in the sentence), the total reading time, and the judgment latency for each sentence were used for analysis along with percentages of errors for each condition. Only data from correct responses were used for analysis. The median values were computed for all the conditions for individual participants and used for statistical analysis. A repeated measure ANOVA with word order (SOV, OSV, and SVO) was performed only with non-anomalous sentences for all the dependent measures. We also performed a repeated measure ANOVA with word order (SOV, OSV, and SVO) and violation (anomalous and non-Anomalous) as factors for all the dependent measures. All the post-hoc analyses in the current study were performed with the Tukey's HSD test.

Reading time of critical verb

Analysis on critical verb reading times for non-anomalous sentences showed a main effect for word order $F(2, 46) = 9.584$, $MSE = 15451.63$, $p < .001$. Planned comparisons between different word orders showed that critical verb reading time was faster for SVO compared to OSV, $t(23) = 2.421$, $p < .05$ and SOV, $t(23) = 3.964$, $p < .001$. In addition, verb reading time for OSV was faster than that of SVO, $t(23) = 2.193$, $p < .05$.

The analysis (see Fig. 1) with both anomalous and non-anomalous sentences showed a significant main effect for word order, $F(2, 46) = 12.939$, $MSE = 33372.01$, $p < .001$. Planned comparisons showed that critical verb reading was significantly faster for SVO compared to SOV, $t(23) = 4.085$, $p < .001$, and OSV $t(23) = 3.196$, $p < .005$. The critical verb reading time was also significantly faster for OSV compared to SOV, $t(23) = 2.33$, $p < .05$. The main effect for anomaly was significant, $F(1, 23) = 13.645$, $MSE = 14105.12$, $p < .001$. The critical verb reading times for the non-anomalous sentences were faster than the anomalous sentences.

In addition, the interaction between word order and anomaly was also found to be significant, $F(2, 46) = 4.768$, $MSE = 7925.505$, $p < .05$, indicating that detection of anomaly is dependent on the word order type. Post-hoc comparisons showed that non-anomalous sentences were significantly faster than anomalous sentences for OSV, $t(23) = 7.034$, $p < .001$. With SOV and SVO, the difference was not significant, although the trends were similar compared to OSV sentences. With respect to anomalous sentences, SVO was faster than OSV, $t(23) = 11.04$, $p < .0001$ and SOV $t(23) = 11.023$, $p < .0001$. The difference between OSV and SOV anomalous sentences was not significant. With respect to non-anomalous sentences, once again SVO was faster than OSV, $t(23) = 4.892$, $p < .05$ and SOV $t(23) = 8.619$, $p < .0001$. The difference between OSV and SOV was close to significance, $t(23) = 3.727$, $p = .128$.

Total Reading Time

The analysis with total reading times for only non-anomalous sentences showed that the effect for word order was not significant, $F(2, 46) = 1.349$, $MSE = 57960.01$, $p = .27$. However, planned comparisons showed a close to significance difference with faster total reading times for SVO compared to SOV, $t(23) = 1.74$, $p = 0.095$. There was no difference in total reading times between SVO and OSV sentences.

The analysis (see Fig. 2) with both anomalous and non-anomalous sentences indicated that the effect of word order was not significant. The main effect of anomaly was significant, $F(1, 23) = 17.554$, $MSE = 45336.31$, $p < .0001$ with longer sentence reading times for anomalous sentences compared to non-anomalous sentences. The interaction between word order and anomaly was close to significance, $F(2, 46) = 2.63$, $MSE = 39438.37$, $p = .08$. Post-hoc comparisons showed that non-anomalous sentences were read faster than anomalous sentences OSV, $t(23) = 5.167$, $p < .05$ and SVO, $t(23) = 4.808$, $p < .05$. The difference between OSV and SOV anomalous sentences was close to significance, $t(23) = 3.793$, $p = .11$.

Judgment Latency

The analysis with total reading times for only non-anomalous sentences showed a close to significant effect for word order, $F(2, 46) = 3.035$, $MSE = 18040.061$, $p = .058$. Planned comparisons showed a significantly faster judgment times for SVO compared to SOV, $t(23) = 2.307$, $p < .05$. There was no difference in total reading times between SVO and OSV sentences.

With judgment latency analysis with both anomalous and non-anomalous sentences (see Fig. 3), the main effect for word order was significant $F(2, 46) = 3.85$, $MSE = 19359.182$, $p < .05$. Planned comparisons showed that judgment latency for SVO was significantly faster than

OSV, $t(23) = 3.842$, $p < .05$. Judgment latency for SOV were not significantly different compared to OSV and SVO. The main effect for violation was not significant. The interaction between word order and violation was significant, $F(2,46) = 3.473$, $MSE = 18521.27$, $p < .05$, indicating that judgment latency for violations is dependent on word order type. Post-hoc comparisons showed that judgment latency for non-anomalous sentences was close to significantly different from anomalous for OSV, $t(23) = 3.846$, $p = .109$. However there were no difference between anomalous and non-anomalous sentences for SOV and SVO. The difference between OSV and SOV anomalous sentences were close to significance, $t(23) = 3.864$, $p = .107$. The rest of the comparisons were not significant.

Judgment Error Rate

The analysis with only non-anomalous sentences showed a significant main effect for word order, $F(2, 46) = 16.185$, $MSE = 29.301$, $p < .0001$. Planned comparisons showed that judgment errors for OSV were lesser compared to SOV, $t(23) = 5.688$, $p < .0001$ and SVO, $t(23) = 3.143$, $p < .01$. Judgment errors for SOV were lesser compared to SVO sentences, $t(23) = 2.593$, $p < .05$.

With the analysis on error rates for anomaly judgments (see Fig. 4) with both anomalous and non-anomalous sentences the main effect for word order was significant $F(2, 46) = 19.806$, $MSE = 32.273$, $p < .001$. Pair-wise comparisons showed that errors for SOV and SVO were significantly higher compared to OSV, $t(23) = 5.854$, $p < .0001$, and $t(23) = 5.55$, $p < .0001$ respectively. There was no significant difference between error rates for SOV and SVO sentences. The main effect for anomaly was not significant.

In addition, the interaction between word order and violation was significant, $F(2, 46) = 9.293$, $MSE = 25.805$, $p < .0001$, indicating that anomaly detection is dependent on the word order type. Post-hoc comparisons showed that error rates for non-anomalous sentences was significantly more than anomalous sentences for SOV word order type, $t(23) = 6.162$, $p < .01$. With OSV and SVO word order type, there were no differences in error rates for anomalous and non-anomalous sentences. With anomalous sentences, there were significantly more errors for SVO compared to SOV, $t(23) = 4.421$, $p < .05$ and OSV, $t(23) = 7.903$, $p < .01$ sentences. With non-anomalous sentences, there were significantly more errors for SOV compared to OSV, $t(23) = 4.421$, $p < .05$. In addition, SVO had significantly more errors than OSV, $t(23) = 4.421$, $p < .05$.. The difference between SVO and SOV non-anomalous sentences was close to significant, $t(23) = 4.152$, $p = .071$.

Discussion

The present study investigated the effect of word order variation on semantic anomaly detection using a self-paced reading paradigm in Hindi. There was a significant effect of word order on verb reading times. Interestingly, verb reading times showed that canonical structures are not necessarily easy to process. With non-anomalous sentences, reading times for verbs were the least for SVO and highest for SOV with OSV in between. The pattern of word order results, especially with non-anomalous sentences strongly suggests a working memory effect arising from distance between the subject and verb. These results support a theoretical position where sentence processing depends on proximity between subject and verb but it is not necessary that such structures be canonical structures for the concerned language. Importantly, the non-canonical SVO structure in Hindi shows a processing advantage over SOV on all dependent measures. Therefore, processing demands are higher when subjects and verbs are not adjacent.

Reading times on the critical verb regions indicate an online effect of the sentential structure that is modulated by the distance between the subject and the verb. Interestingly most of the studies reviewed in this paper utilizing the self-paced reading paradigm have not used reading times on critical sentential constituents i.e. verbs as dependent measures. Results from our study on this measure are not comparable to other studies except measures like total reading times and judgment latency. Our results show important online effects of word order variation on sentence comprehension.

Current accounts assume that sentence processing is incremental and probabilistic (Altmann & Mirkovic, 2009). In fixed word order languages like English, language comprehenders have less ambiguity in terms of word order due to lack of scrambling. However, with free word order languages, language comprehenders anticipate multiple sentential patterning due to the fact that different word orders are possible. When the first noun phrase of a sentence in Hindi is identified as subject, the next constituent could be the verb or could be another noun phrase. In inflectional languages like Hindi, case markers on the subject i.e. ergative or no marker often help disambiguate the confusion about the identity of the next upcoming structure (Choudhary, Schlesewsky, Roehm & Bornkessel-Schlesewsky, 2009). If the subject has been identified and it does not contain an ergative marker, then language comprehenders expect a verb. In such situations, the proximity of the verb to the subject facilitates sentence processing as opposed to other structures. While subject and verb are in proximity in OSV sentences as well, the object initial processing might create difficulty. Ambiguities associated with the initial noun in object initial structures can lead to multiple interpretations especially with OSV sentences that would impose additional processing demands on sentence processing (Choudhary, 2010; Mishra, 2007). Choudhary (2010) reported larger processing costs in the case of core arguments in which the object appeared before the subject.

With respect to anomaly detection, word order interacted with anomaly, suggesting that even in a flexible word order language like Hindi, different word order types are processed differently with respect to anomaly detection. Canonicity of the sentence does not necessarily facilitate speed or accuracy of anomaly detection. Similar to non-anomalous sentences, reading time for critical verb was faster for SVO compared to OSV and SOV sentences for anomalous sentences. There was no difference in critical verb reading times between SOV and OSV for anomalous sentences since anomaly detection is modulated by the adjunct clause. The results with anomalous sentences also indicate a role of working memory in anomaly detection.

Judgment latency potentially reflects costs associated with reanalysis involving verification of syntactic and semantic information that has been read i.e., after all the words in the sentence have been read and lexical integration has taken place. This measure reflects an offline activity as opposed to reading times on critical words like the verb. Reading times reflect an online measure and therefore are indicative of the processing difficulties faced by the parser while the material is being comprehended. Accuracy is an offline measure that involves decision processes after the sentence has been read. Our results show a significant effect of sentence type on judgment latency. Subjects took the least time to judge correctly a SVO sentence. This finding is also dissimilar to those studies that indicated minimal times for the canonical structures. Here again, it could be the case that locality, the proximity of the subject and the verb resulted in better processing and needed less working memory resources compared to a SOV or a OSV structure. The participants may have detected the anomaly much before the ending with SVO sentences, which may have been delayed for other types of sentence due to larger demand in working memory resources (Bahlmann, 2007). In case of the OVS structure, even though the subject and the verb are together, the object initial processing makes it difficult as shown in earlier studies on object-first processing (Mishra, 2007) in spite of adjacency of subject and verb.

The effect of anomaly itself is reversed for the canonical SOV sentences with faster judgment times and fewer errors for anomalous sentences compared to non-anomalous sentences. This is different from SVO and OSV sentences that actually show generally poorer or same performance with anomalous compared to non-anomalous sentences. While there is no overall advantage for the canonical sentences, it appears that judging anomaly is easier than judging normality with such sentences.

Overall reading times also indicate that there is no advantage for canonical sentences and the trends were similar to the other measures with SVO being the best condition, even though this was not significant. The results provide no support for a canonicity or syntactic "filler gap" model of sentence processing, and raise important questions regarding the effects of scrambling in Hindi. Thus, our results on reading times and judgment latency in the context of semantic anomaly judgments are best explained using a working memory-based or resources-based locality account, in which processing is modulated by the distance between the critical sentential elements. These results bring another dimension to the existing conflict in findings i.e. preferences for a canonical word order in Basque (Erdocia, et al., 2009), German (Weyerts, et al., 2002) and Japanese (Kim et al., 2009) as opposed to the null effect in other Japanese studies (Tamaoka, et al., 2003).

German is a flexible word order language like Hindi and has typological similarities like agreement and gender marking. Using self-paced reading, Weyters et al., (2002) found that SVO sentences are easier to process than SOV sentences. However, SVO happens to be the canonical word order type for German where as for Hindi it is SOV. Rosler et al., (1998) found faster comprehension times with S-IO-DO compared to S-DO-IO. The object-initial sentences took longer to comprehend than subject-initial sentences. The results with German have been interpreted as preference for the canonical structures in processing. However, the results with German are also consistent with a working memory based explanation that would also predict better performance with SVO in comparison with SOV or OSV sentences. The results with Hindi are similar to those with German and both can be explained using a locality-dependent working memory based explanation of sentence processing.

The results with Hindi are not entirely consistent with results from other languages. For example, the results from Basque (Erdocia, et al., 2009) indicate an advantage for canonical structures (SOV vs. OSV). One study with Japanese did not show any effect for word order (Tamaoka, et al., 2003) but another study found an effect for word order with better performance with SOV compared to OSV sentences (Kim et al., 2009). It is to be noted that SVO (the best condition in our study) was not used in the studies on Basque and Japanese (Erdocia et al., 2009; Kim et al., 2009) and it would be interesting to see whether the advantage for canonical structure would still be present with the addition of SVO sentences. Clearly, more work on other free word order languages is required to substantiate the findings from Hindi. In summation, it can be argued that canonicity does not automatically provide an advantage at least with semantic anomaly detection in sentences. Therefore, understanding the actual effect of scrambling may require further explorations into the specific typological features of a given language and a re-examination of the very concept of “canonicity” or “scrambling complexity”.

Many previous studies have explored the effect of word order variations or scrambling on sentence comprehension in different languages using both behavioral as well as cognitive neuroscience methods. Models like syntactic prediction locality theory and trace reactivation account predict difficulty in processing the non-canonical structure. However, results of this study do not support such positions. Instead, our results indicate a critical role of working memory on semantic anomaly detection in Hindi. Further studies are needed to elucidate the critical role of syntactic structure and other cognitive processes like working memory on semantic anomaly detection and more generally on Hindi language comprehension.

Implications for Acquisition

Very few studies have directly looked into the mechanisms of word order acquisition in free word order languages. Often the lack of a model of sentence processing with different word orders in such languages makes experimental results difficult to interpret. However, theories of word order processing do have implications for child language acquisition. As in Hindi, learners encounter scrambling in the linguistic input much early compared to English. This means, learners of a free word order language, must consider the various versions of the sentences to be similar in meaning with surface structural differences. From a language acquisition point of view, very early parameter setting hypothesis of Wexler (1998) postulates that core aspects of sentence features such as position of verbs with respect to objects and the presence and absence of null subjects are acquired first. Earliest settings of basic parameters may allow the learner of a free word order language to follow the canonical pattern of that language. Identification of other patterns depend on their difference with respect to the canonical patters. For example, in a recent study in Japanese, a free word order language, Sugisaki (2008) observed emergence of the basic word order first in young Japanese child's utterances. However, currently there is no such data in Hindi to support this conclusion. Therefore, what processing advantages the canonical or non-canonical word order has for language learning in free word order languages remains an open empirical question.

A possible advantage of processing a free word order language could be flexibility in cognitive processing. This flexibility in processing may arise because of processing diverse types of scrambled structures in a free word order language as opposed to a rigid word order language. This suggests that, Hindi children may demonstrate higher cognitive flexibility (e.g. faster development of working memory capacity) tested on some other non-linguistic tasks compared to age matched children speaking a strict word order language. Better working memory capacity on non linguistic tasks have been found to be predictive of efficient text integration and have been considered a suitable measure to differentiate good from poor readers (Carretti, Borella, Cornaldi, & De Beni, 2009; Nation, Adams, Bowyer-Crane, & Snowling, 1999**)**. Performance on working memory tasks has also been shown to predict children's performance on language-based tasks such as reading comprehension and inference making (Cain, Oakhill, & Bryant, 2004). The role played by working memory in sentence processing and the presence of different word orders might benefit other cognitive processes including the acquisition of language and literacy. Future studies need to look into the effects of sentence processing on language acquisition and literacy.

References

Altmann, G. T. M., & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science, 33*, 1–27.

Bahlmann, J., Rodriguez-Fornells, A., Rotte, M., & Munte, T. F. (2007). An fMRI study of canonical and noncanonical word order in German. *Human Brain Mapping, 28*, 940-949.

Braze, D., Shankweiler, D., Ni, W., & Palumbo, L. C. (2002). Readers' eye movements distinguish anomalies of form and content. *Journal of Psycholinguistic Research, 31,* 25–44.

Carretti, B., Borella, E., Cornoldi, C., & De Beni, R. (2009). Role of working memory in explaining the performance of individuals with specific reading comprehension difficulties: A meta-analysis. *Learning and Individual Differences*, *19*, 246-251.

Cain, K., Oakhill, J., & Bryant, P. (2004). Children's reading comprehension ability: Concurrent prediction by working memory, verbal ability, and component skills. *Journal of Educational Psychology*, *96*, 31-42.

Choudhary, K. K. (2010). Incremental argument interpretation in a split ergative language: electrophysiological evidence from Hindi. Ph.D. Thesis submitted to University of Leipzig, Germany to be published in MPI Series of Cognitive Sciences, Leipzig.

Choudhary, K. K., Schlesewsky, M., Roehm, D., & Bornkessel-Schlesewsky, I. (2009). The N400 as a correlate of interpretively relevant linguistic rules: Evidence from Hindi. *Neuropsychologia, 47*, 3012-3022.

Chomsky, N. (1965). *Aspects of theory of syntax*. MA: MIT Press.

Chomsky, N. (1981). *Lectures on government and binding*. Foris: Dordrecht.

Comrie, B. (1981). *Language Universals & Linguistic Typology*. Chicago: The University of Chicago Press.

Dayal, V., & Mahajan, A. (2004). *Clause Structure in South Asian Languages*. Kluwer Academic Publishers.

de Villiers, P. A., & de Villiers, J. G. (1972). Early judgments of semantic and syntactic acceptability by children. *Journal of Psycholinguistic Research*, *1*, 299-310.

Erdocia, K., Laka, I., Mestres-Misse, A., Rodriguez-Fornells, A. (2009). Syntactic complexity and ambiguity resolution in a free word order language: Behavioral and electrophysiological evidences from Basque. *Brain and Language, 109*, 1-17.

Fodor, J. D. (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry, 9*, 427-473.

Frazier, L. (1995). Constraint satisfaction as a theory of sentence processing. *Journal of Psycholinguistic Research, 6,* 437-468.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition, 6,* 291–325.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition*, *68*, 1-76.

Karimi, S. (Ed.) (2003). *Word order and scrambling.* Blackwell Publishing.

Kayne, R. S. (1994). *The antisymmetry of syntax*. Cambridge, MA: MIT Press.

Kim, J., Koizumi, M., Ikuta, N. et al., (2009). Scrambling effects on the processing of Japanese sentences: An fMRI study. *Journal of Neurolinguistics, 22,* 151-166.

Kolk, H., & Weijts, M. (1996). Judgment of semantic anomaly in aggrammatic patients. *Brain and Language*, *54*, 86-135.

Kuperberg, G., Caplan, D., Sitnikova, T., Eddy, M., & Holcomb, P. (2006). Neural correlates of processing syntactic, semantic, and thematic relationship. *Language and Cognitive Processes*, *21*, 489-530.

Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic anomalies. *Nature, 307,* 161-163.

Lewis, R. L., & Vasishth, S. (2005). An activation model of sentence processing as skilled memory retrieval. *Cognitive Science, 29,* 375-419.

Mishra, R. K. (2007). Spatial premise integration in Hindi. *Indian Journal of Applied Linguistics, 33,* 103-118.

Miyagawa, S. (2005). EPP and semantically-vacuous scrambling. In J. Sabel and M. Saito (Eds.), *The Free Word Order Phenomenon: Its Syntactic Sources and Diversity* (181-220), Berlin: Mouton de Gruyter.

Mohanan, K. P. & Mohanan, T. (1994). Issues in word order. In M. Butt, T. H. King, & G. Ramchand (Eds.), *Perspectives on Word Order in South Asian Languages*, Stanford: CSLI Publications.

Nation, K., Adams, J. W., Bower-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology*, *73*, 139-158.

Nakayama, M. (1995). Scrambling and probe recognition. In R. Mazuka & N. Nagai (Eds.), *Japanese sentence processing*. Hillsdale, HJ: Erlbaum.

Nakayama, M. (1999). Sentence processing. In M. Tsujimura (Ed.), *The Handbook of Japanese Linguistics* (pp. 398–424), Malden, MA: Blackwell.

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research, 18*, 5-19.

Ni, W., Fodor, J. D., Crain, S., & Shankweiler, D. (1998). Anomaly detection: Eye movement patterns. *Journal of Psycholinguistic Research, 27,* 515–540.

Pechmann, T., Uszkoreit, H., Engelkamp, J., & Zerbst, D. (1994). *Word order in the German middle field: Linguistic theory and psycholinguistic evidence* (CLAUS Report No. 43). Department of Computational Linguistics, Saarland University.

Rosler, F., Pechmann, T., Streb, J., Roder, B., & Hennighausen, E. (1998). Parsing of sentences in a language with varying word order: Word-by-word variations of processing demands are revealed by event-related brain potentials. *Journal of Memory and Language, 38,* 150-176.

Sugisaki, K. (2008). Early acquisition of basic word order in Japanese. *Language Acquisition, 15*, 183-191.

Schlesewsky, M., Bornkessel, I., & Frisch, S. (2003). The neurophysiological basis of word order variations in German. *Brain and Language, 86,* 116-128.

Sekerina, I. (1997). The syntax and processing of Russian scrambled constructions. Unpublished Ph.D Dissertation. City University of New York: NY.

Sekerina, I. (2003). Scrambling processing: Dependencies, complexity, and constraints. In S. Karimi (Ed.), *Word Order and Scrambling* (pp. 301–324), UK: Blackwell.

Tamaoka, K., Sakai, H., Kawahara, J., & Miyoaka, Y. (2003). The effects of phrase-length order and scrambling in the processing of visually presented Japanese sentences. *Journal of Psycholinguistics Research, 32*, 431-454.

Wexler, K. (1998). Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua, 106*, 23-79.

Weyerts, H., Penke, M., Munte, T. F., Heinze, H. J., & Clahsen, H. (2002). Word order in sentence processing: An experimental study of verb placement in German. *Journal of Psycholinguistics Research, 31,* 211-268.

Wolff, S., Schlesewsky, M., Hirotani, M., & Bornkessel-Schlesewsky, I. (2008). The neural mechanisms of word order processing revisited: Electrophysiological evidence from Japanese. *Brain and Language, 107,* 133-157.

Yamashita, H. (1997). The effects of word-order and case marking information on the processing of Japanese. *Journal of Psycholinguistic Research, 26,* 163–188.

Figure captions

Figure 1: Verb reading times for anomalous and non-anomalous sentences as a function of word order.

Figure 2: Total reading times for anomalous and non-anomalous sentences as a function of word order.

Figure 3: Judgment latencies for anomalous and non-anomalous sentences as a function of word order.

Figure 4: Error rates for anomalous and non-anomalous sentences as a function of word order.

Appendix

SOV anomalous

1. Badal pani bharne se chaunke
   Cloud rain because of shocked
   Cloud shocked when it rained
2. Medhak talab sukhne se pighale
   Frogs pond dry because of melted
   Frogs melted when pond dried
3. Chappal chalne se thaki
   Sleeper walk because of tired
   Sleeper got tired when walked
4. Kursi pani padne se uchli
   Chair water drop because of jumped
   Chair jumped when water dropped
5. Suraj sham hone par bhaga
   Sun evening because of ran
   Sun ran in the evening
6. Kanch patthar par lagne se roya
   Glass stone on hit because of cried
   Glass cried when got hit with the stone
7. Dhuna ghar ke jalne se barsa
   Smoke house of burn because of rained
   Smoke rained when house got burnt
8. Pankha bijli jane se gaya
   Fan current went because of sang
   Fan sang when current went off
9. Patte hawa chalne se hanse
   Leaves wind blow because of laughed
   Leaves laughed when wind blew
10. Paudha pani padne se uda
    Plant water because of flew
    Plant flew when watered

OSV anomalous

1. Patthar par rakhne se cycle boli
   Stone on keep because of cycle spoke
   Cycle spoke when kept on stone.
2. Bahar rakhne se kitab nachi
   Outside keep because of book danced
   Book danced when kept outside
3. Mez par khula rakhne se pen chala
   Table on open keep because of pen walked
   Pen walked when kept open on table
4. Lakadi ke jalane se dhnua phata
   Wood burn because of smoke torn
   Smoke torn by burning wood
5. Uske marne se kutta uda

He beat because of dog flew
Dog flew when he beat

6. Bahar rakhne se botal jagi
Outside keep because of bottle woke
Bottle woke up when kept outside
7. Andhi ane se ped jhagade
Storm came because of trees quarreled
Big trees quarreled when it was storming
8. Tar tutne se pankha hansa
Wire break off because of fan laughed
Fan laughed when wire broke off
9. Talab sukhne se machli tahali
Pond dried because of fish walked
Fish walked when pond dried
10. Sher ko dekhkar hiran gurraya
Lion see because of deer roared
Deer roared when saw a loin

SVO anomalous

1. Ped chillaye kulhadi se katne par
Tree cried hammer because of cut
Tree cried when cut with hammer
2. Kapda soya dhup me rakhne se
Cloth slept sunshine on keep because of
Cloth slept when kept in sunshine
3. Cycle royi zameen par girne se
Bicycle cried ground on fall because of
Bicycle cried when fell on the ground
4. Kabutar taira pinjara khulne se
Pigeon swam cage open because of
Pigeon swam when cage opened
5. Pen nacha zameen par girne se
Pen danced ground on fall because of
Pen danced when fell on the ground
6. Dhuna bhaga lakdi ke jalne se
Smoke ran wood of burn because of
Smoke ran when wood got burnt
7. Kitab soyi mez par rakhne se
Book slept table on keep because of
Book slept when kept on the table
8. Imarat gayi hawa ke chalne se
Building sang wind of blew because of
Building sang when wind blew
9. Phul bola tod lene se
Flower spoke pluck because of
Flower spoke when plucked
10. Ghar dauda aag ke lagne se

House ran fire of burn because of
House ran when on fire

SOV non-anomalous

1. Pankha taar tutne se gira
   Fan wire break because of fell
   Fan fell when the wire broke
2. Ghar aag lagne se jala
   House fire aux because of burned
   House burned in fire
3. Danda lohe par patakne se tuta
   Stick iron on hit because of broke
   Stick broke when hit with iron
4. Kapda pani barasne se bheega
   Cloth water fall because of got wet
   Cloth got wet when water fell
5. Admi saanp ke katne se mara
   Man snake by bite because of died
   Man dies because of snake bite
6. Kagaz hawa chalne se ude
   Paper air move because of flew
   Paper flew in air
7. Kursi putai karne se chamki
   Chair polish aux because of shone
   Chair shone when polished
8. Hathauda nadi me girne se khoya
   Hammar river in drop because of lost
   Hammar was lost when dropped in the river
9. Khargosh bade janwar ko dekhkar bhaga
   Rabbit big animal see because of ran
   Rabbit ran when it saw a big animal
10. Anda zameen par girne se phuta
    Egg ground on fall because of broke
    Egg broke when fell on ground

OSV non-anomalous

1. Patthar marne se kanch tuta
   Stone hit because of glass broke
   Glass broke when hit with stone
2. Aag lagne se kapda jala
   Fire burn because of cloth burnt
   Cloth burnt in fire
3. Kulhadi marne se darwaza tuta
   Axe hit because of door broke
   Door broke when hit with axe
4. Badh aane se pul tuta
   Flood came because of bridge fell
   Bridge fell when flood came

5. Patthar par patakne se nariyal tuta
   Stone on hit because of coconut broke
   Coconut broke when hit with stone
6. Aandhi ane se ped gira
   Storm come because of tree fell
   Tree fell when storm came
7. Pani barasne se Kapda bheega
   Water fell because of Cloth wet
   Cloth got wet when water fell
8. Saanp ke katne se Admi mara
   Snake of bite because of Man died
   Man died because of snake bite
9. Zameen par girne se Anda phuta
   Ground on fell because of egg smashed
   Egg got smashed when fell on ground
10. Pani padne se kitab bheegi
    Water fell because of Book wet
    Book got wet in water

SVO non-anomalous

1. Kaanch pighla tapman badhane se
   Glass melted temperature rise because of
   Glass melted when temperature rose
2. Tala tuta hathaudi marne se
   Lock broke hammar hit because of
   Lock broke when hit with hammar
3. Kagaz jala aag me dalne se
   Paper burned fire in throw because of
   Paper burned when thrown in fire
4. Sher gurraya shikari ko dekhkar
   Lion roared hunter see because of
   Lion roared when it saw the hunter
5. Pani gira botal tutne se
   Water fell bottle broke because of
   Water fell when bottle broke
6. Billi bhagi kutte ke aane se
   Cat ran dog came because of
   Cat ran when the dog came
7. Saabun gala pani padne se
   Soap melted water fall becaue of
   Soap melted in water
8. Khidki tuti kulhadi marne se
   Window broke hammar hit because of
   Window broke when hit by hammar
9. Paudha jhulsa dhoop lagne se
   Plant dried sun because of
   Plant dried because of sun

10. Ghar ujda tuufan ane se
House demolished storm come because of
House demolished in the storm

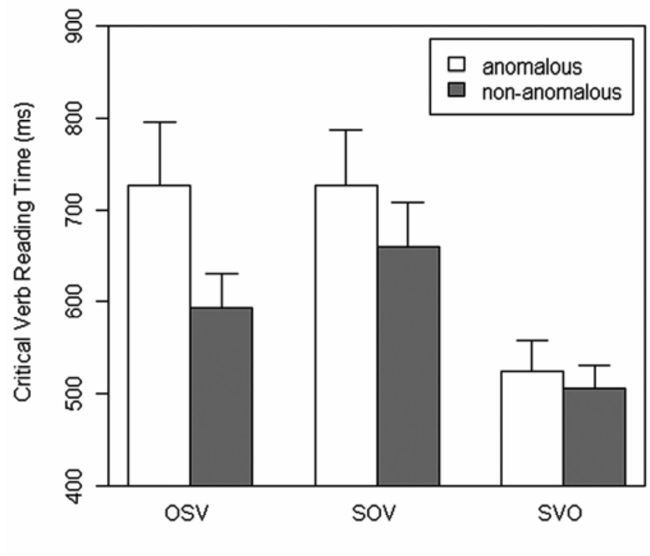

anomalous
non-anomalous
Critical Verb Reading Time (ms)
900
800
700
600
500
400
OSV
SOV
SVO

Figure 1

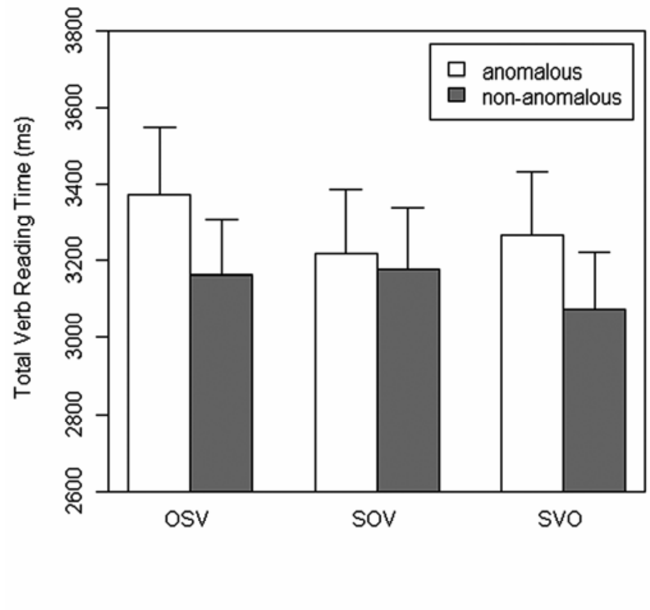
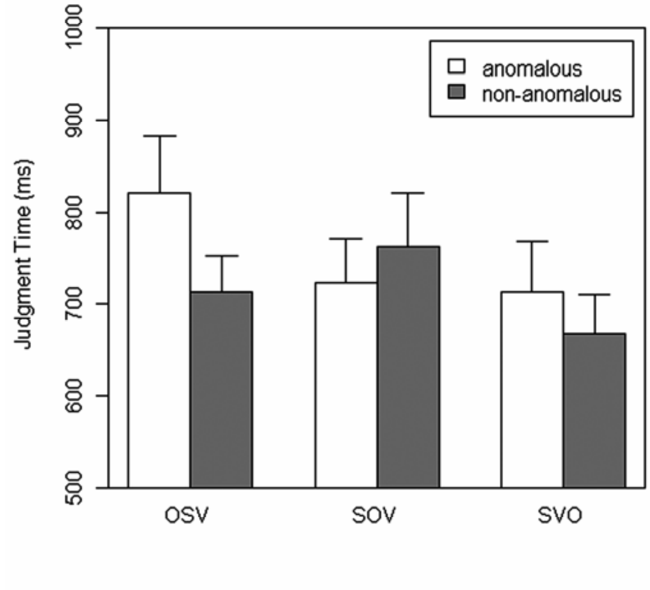

Total Verb Reading Time (ms)
3800
3600
3400
3200
3000
2800
2600
anomalous
non-anomalous
OSV
SOV
SVO

Figure 2



1000
900
800
700
600
500
Judgment Time (ms)
anomalous
non-anomalous
OSV
SOV
SVO

Figure 3

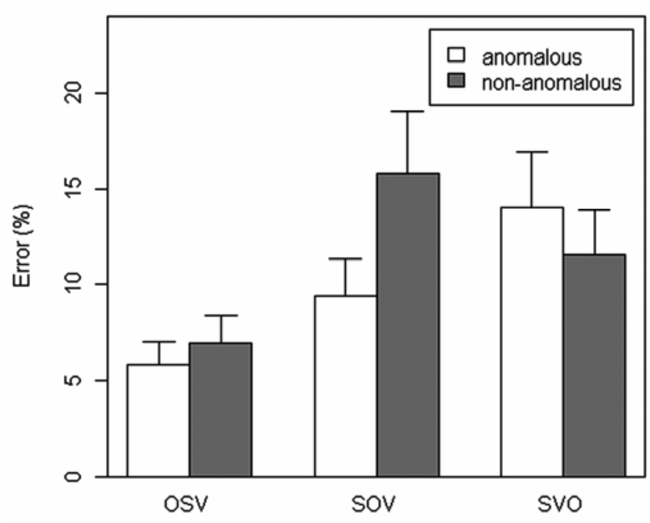


Figure 4