Title
Synthetic Brain Imaging of English Past Tense Inflection

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 31(31)

ISSN
1069-7977

Authors
Ruh, Nicolas
Westermann, Gert

Publication Date
2009

Peer reviewed

Synthetic Brain Imaging of English Past Tense Inflection

Gert Westermann (gwestermann@brookes.ac.uk)
Nicolas Ruh (nruh@brookes.ac.uk)

Department of Psychology, Oxford Brookes University
Oxford OX3 0BP, UK

Abstract

Brain imaging studies of English past tense inflection have found dissociations between regular and irregular verbs, but no coherent picture has emerged to explain how these dissociations arise. Here we use synthetic brain imaging on a neural network model to provide a mechanistic account of how regional dissociations between regular and irregular processing in adults can emerge in a single mechanism system with experience-dependent structural development. We show that these dissociations arise from a combination of different statistical properties of verbs relating to frequency, relationships to other verbs, and phonological complexity. The model generates predictions about the patterning of active brain regions for different verbs that can be tested in future brain imaging studies.

Keywords: English past tense; verb inflection; connectionist modeling, synthetic brain imaging, neuroconstructivism

Introduction

The question of how verb inflections, particularly the English past tense, are represented and processed in the brain has been the subject of intense debate over the past 20 years. This is because the past tense contains regular forms (e.g., *look-looked*) that seem to obey a linguistic mental rule, and irregular forms (e.g., *see-saw*) that show properties of associative storage. One view, the Dual-Mechanism or Words-and-Rules Theory (e.g., Pinker, 1999; Pinker & Ullman, 2002) holds that the apparent processing differences between regular and irregular forms are indeed caused by qualitatively different underlying mechanisms: a rule for regulars, and associative storage for irregulars. Another view, closely linked to connectionist approaches to cognitive processing, argues that all forms are processed in a single associative system and that apparent dissociations between regulars and irregulars emerge on the basis of different statistical properties of verbs that affect their ease of processing (e.g., frequency, phonological complexity, number of similar verbs with a similar past tense (*friends*, e.g., *sing* and *ring*), number of similar verbs with a different past tense (*enemies*, e.g., *sing* and *bring*)) or on the basis of different reliance on semantic vs. phonological factors (Joanisse & Seidenberg, 1999; McClelland & Patterson, 2002; Seidenberg & Arnoldussen, 2003; Westermann, 1998). A large amount of empirical and computational work has aimed to provide evidence for either view. One methodology that has been used in several studies is brain imaging with positron emission topography (PET), functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG) and evoked response potentials (ERP). These studies have found differences in brain activation patterns when participants inflect regular and irregular verbs, respectively, leading some researchers to claim that these data provide support for a dual mechanism system, with the rule component and the associative mental lexicon located in different brain regions (Dhond et al., 2003; Jaeger et al., 1996; Lavric et al., 2001). For example, in a seminal study by Jaeger et al. (1996) using PET, participants were asked to generate past tense forms of visually presented monosyllabic verb stems. Results showed that although many brain regions were activated by all verbs, production of regulars selectively activated left dorsolateral prefrontal cortex and left anterior cingulate cortex. Irregulars led to higher overall activation and involved occipital visual processing areas. These systematic differences between both verb types were interpreted by the authors as strong evidence for the dual-mechanism account of inflection.

In another study, Dhond et al. (2003) asked participants to covertly generate past tense forms of visually presented verb stems while imaging their brains using MEG. Dhond et al. also found that generation of regulars and irregulars activated many areas in common, but that processing of regulars led to greater activation in left inferior prefrontal areas (Broca's area), and processing of irregulars preferentially activated left occipitotemporal cortex as well as right dorsolateral prefrontal cortex. These results were interpreted as indicating that regulars activated rule-based grammar regions and irregulars activated areas involved in associative retrieval of forms, corresponding directly to the dual-mechanism theory.

However, the results of these and other studies have been controversial. One problem is that specific methodological choices can strongly affect results. For example, because of the low temporal resolution of PET, Jaeger et al. (1996) used a block design in which all regular verbs and all irregular verbs were presented together. However, this design introduces the confound that participants could develop strategies for regular but not for irregular verbs, suggesting that differences between both verb types should be found independently of the nature of the underlying processing mechanisms (Seidenberg & Hoeffner, 1998).

Several more recent imaging studies have investigated the possibility that observed activation differences between regular and irregular verbs are due to statistical factors of verbs and not to separate underlying mechanisms. These studies have attempted to carefully control a number of statistical factors within their experimental stimuli. In an fMRI study of past tense production in which verbs were matched for past tense log frequency, friend-enemy ratio, stem letter-length, and stem and past tense syllable-length, Desai et al. (2006) found widespread overlapping activation for all verbs. Greater activation for irregulars compared to regulars was found in several regions in inferior frontal, precentral and parietal cortex bilaterally. Greater activation for regulars was found in the left dorsal superior temporal gyrus, involving the primary auditory areas and the planum temporale. For a subset of these verbs that were matched for phonological complexity of the past tense form no regions were activated more for regulars than for irregulars. Desai et al. explained the widespread activation of brain regions for irregular verbs with higher demands on attention, working memory and response selection for generating the past tense forms of these verbs. The fact that both regular and irregular production activated Broca's area was seen as contradicting the dual-mechanism account which assumes that regular, but nor irregular forms are generated through a mental grammar instantiated in Broca's area (Ullman et al., 1997). Greater activation in auditory areas for regulars was explained with regular forms being phonologically more complex than irregular forms. Therefore, despite double dissociations between regular and irregular verbs these results were interpreted as evidence for a single-mechanism view of inflection processing.

In a similar study Joanisse & Seidenberg (2005) presented fMRI data from covert past tense production experiments. Like all previous studies they found that production of regulars and irregulars activated common areas in both hemispheres. Regulars, as well as irregulars that are phonologically similar to regulars (e.g., *burnt*, *slept*), additionally activated the inferior frontal gyrus bilaterally. Overall, irregulars did not activate any area more than regulars. These results were also interpreted as conforming to a single-mechanism view of inflection processing.

In summary, previous imaging studies, despite each reporting dissociations between regular and irregular verbs, have yielded inconclusive results. Not only differed the activated regions for specific verb types greatly between studies, but one study (Joanisse & Seidenberg, 2005) reported activation of distinct brain regions for regulars but not irregulars, another (Desai et al., 2006) reported the opposite result with distinct regions active for irregulars but not for regulars when verbs were matched phonologically, and two other studies (Dhond et al., 2003; Jaeger et al., 1996) reported a double dissociation with some regions more active for regulars and others more active for irregulars. One possible explanation for this inconsistency of results is that indeed statistical factors and not grammatical class determine how a verb is processed, and that these factors differed between the specific verbs used in the described studies. In each study regular and irregular verbs were matched on certain factors, but the choice of factors had little theoretical foundation and differed greatly between studies. Jaeger et al. (1996) matched stem and past tense frequencies (albeit based on a word list that did not distinguish between nouns and verbs and therefore overestimated regular stem frequencies), Lavric et al. (2001) and Dhond et al. (2003) matched for word frequency and letter-length, and Joanisse & Seidenberg (2005) matched for log past tense frequency, imageability, and concreteness. The most careful matching was done in Desai et al.'s (2006) study where past tense log frequency, friend-enemy ratio, stem letter-length, and stem and past tense syllable-length were taken into account, and a sub-group was further matched on the number of phonemes and past tense syllable structure. However, which of these factors affect processing in which way remains an open question. It is therefore also unclear how a processing system that is sensitive to the statistical properties of verbs can give rise to the observed dissociations in the localization of active brain regions.

One approach to answering these questions is to consider how the adult language processing system is shaped through development. Adult psycholinguistics traditionally pays little heed to the mechanisms of language development although a better understanding of developmental trajectories could inform the nature of the adult processing system. In this paper we use computational modeling to ask how the adult inflection processing system emerges through interactions between experience-dependent brain development and experience with verbs that have specific statistical properties. We then use synthetic brain imaging to show that such a system displays visible processing differences between regular and irregular verbs without relying on built-in dissociable processing modules. Finally, we investigate which statistical properties account for the observed dissociations, generating predictions for behavioral and imaging studies.

The computational model described here is based on the idea of neuroconstructivist development. The theory of neuroconstructivism (Mareschal et al., 2007; Westermann et al., 2007) sees cognitive development as a trajectory shaped by intrinsic (genes, mechanism of experience-dependent brain development, principles of neural processing) and extrinsic (structured environment, social interactions) constraints. The adult state is a point (or more precisely, a region) on this trajectory. Based on this view, in principle a single processing theory should account for development and adult processing. This implies that in order to understand the adult system it is important to understand the developmental process that has given rise to this system.

The model described here follows the neuroconstructivist approach and integrates structural changes that mimic, on an abstract level, the experience-dependent structural development of cortical regions, allowing for the adaptation of neural circuits to a specific learning task. This model aims to provide a mechanistic account of how a single-mechanism processing system that is sensitive to statistical properties of learned verbs gives rise to dissociations between regular and irregular verbs through a neuroconstructivist developmental process.

Synthetic Brain Imaging

Synthetic brain imaging (SBI, Arbib et al., 2000) applies the idea of brain imaging – to compare brain region activation profiles between different conditions to gain insights into underlying processing mechanisms – to artificial neural networks. In a structured neural network different stimuli will generate activation patterns in different network components and, like in brain imaging, these patterns can be compared between conditions. Although SBI is still in initial stages of exploration, several results have been reported. Previous applications of this method have shown differential activation for nouns and verbs in evolved agent-based networks (Cangelosi & Parisi, 2004). In a model of sentence comprehension (Just, Carpenter, & Varma, 1999) SBI accounted for fMRI data on brain regions involved in processing sentences of different complexities.

While in most cases the structure of an artificial neural network cannot be mapped onto brain structure without unrealistic assumptions about the match between connectionist models and brain development and processing, using structured neural networks to generate synthetic brain imaging nevertheless can be highly useful for theory building and generating predictions for real brain imaging. First, in neural networks, the experimenter has full control over the studied process. The large number of brain areas activated in inflection production tasks suggests that, even when inflecting a stem is contrasted directly with reading the stem, it is difficult to find a baseline condition that differs from the experimental condition only in the inflection process. In a model of verb inflection that takes a verb stem as input and produces its past tense as output, there is nothing but this process. Second, a model allows for the precise analysis of what factors affect differential activation of network components in a much larger set of verbs than those typically used in experiments. Third, the language experience that has shaped a model to reach its final structure is precisely known. This allows for a more precise measurement of the statistical factors that affect processing.

The Model

A prototype of the model described here has previously been used to model acquisition and adult processing of the English past tense (Westermann, 1998, 2000) and the German participle (Ruh & Westermann, 2008) where it displayed a realistic acquisition profile as well as a pattern of non-word generalization that was comparable to adults. The model incorporates the idea that a theoretical account of past tense processing should encompass both acquisition and adult performance. The model is based on the idea of experience-dependent brain development, incorporating constructive and regressive structural changes in the learning process. The model starts out with a minimal architecture (Fig. 1) with predominantly direct connections between input and output. Hidden layer units have a Gaussian rather than the more common sigmoidal activation function, thus forming a *receptive field* (rf) for a region of the input space. All input vectors are positioned in this space according to their phonetic feature values (see below) and a hidden unit will become maximally activated if its position (i.e., the center of its rf) coincides closely with that of the current input. For each input pattern, only the most active hidden unit contributes to the model's output.

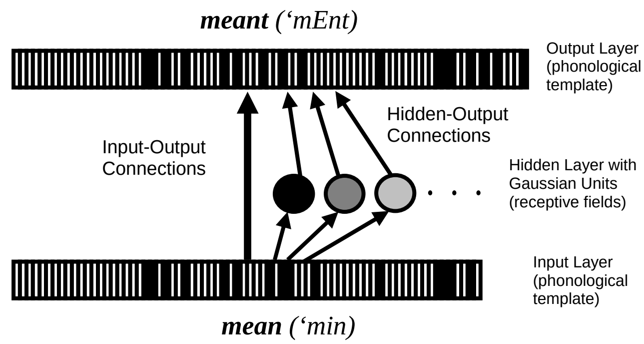

Figure 1: The past tense model

The task of the model was to generate the past tense form of a verb on the basis of its stem. The model learned through weight adaptation after each presented pattern (sweep). The direct connections from the input to the output layer and the connections from the developing hidden layer to the output layer were adapted with the perceptron learning rule (learning rates 0.005 and 0.05, respectively). The positions of the hidden layer receptive fields were adjusted so that the receptive field of a unit moved a fraction towards the position of the current input (learning rate 0.001). The size of the most active hidden unit receptive field was also adjusted depending on the distance between its center and the current input, leading to an increase in size when the receptive field was activated by inputs that fell outside its center. Each hidden unit kept a running counter of its contribution to the out output error.

Structural change in the model occurred when the current structure no longer allowed for improvement in performance. When the averaged network error over 10,000 sweeps was no lower than for the average 30,000 sweeps previously, three new hidden unit receptive fields were inserted at random positions around the existing hidden unit with the highest contribution to output error, and their weights to the output layer were initialized randomly. In this process, a hidden unit whose activation leads to a high output error will become the preferred location for the insertion of new units. Because a high local error is usually caused by one hidden unit being responsible for too many input patterns with conflicting input-output transformations, the insertion of additional resources led to a more fine-grained covering of the input space in those areas where similar sounding verbs have different past tense forms. Regressive events in the model were implemented by pruning hidden units that were not activated for 30,000 sweeps.

This experience-dependent structural development led to a final network structure that was adapted to the specific task of processing past tense inflections of a set of verbs with the given statistical distribution of properties experienced during learning. The assumption here is that the brain likewise adapts to process the past tense on the basis of the specific experiences with verbs.

Corpus

For the training corpus, all mono- and bisyllabic verbs with an unambiguous past tense inflection and a past tense frequency of at least 1 per million entries were extracted from the English part of the CELEX database (Baayen, Piepenbrock, & van Rijn, 1993). The resulting set had 1,271 verb types out of which 111 (8.73% types, 46.00% tokens) were irregular. During training, for every sweep a new verb was drawn from this corpus on the basis of past tense frequencies.

Verbs were represented in the following way: each syllable was fitted into a slotted xCCCVCC template with consonants aligned to the outer boundaries of the syllable to generate a maximum overlap for the initial and final consonants. A single slot (x) was used to indicate the presence or absence of stress on this syllable. Individual phonemes were encoded by phonetic feature vectors, following the binary version of the PatPho coding scheme (Li & MacWhinney, 2002) which requires 6 features per vowel and 7 features per consonant. The presence or absence of a feature was encoded by a value of 1 or -1, respectively, and all features for an empty phoneme slot were set to 0. The stem of a verb was encoded by 84 bits. The past tense form had an additional VC suffix (13 bits).

Performance

Ten different networks were trained for 2m sweeps each. All models reached 100% correct performance, on average after 1,247,800 sweeps (s.d.=272). The average final number of hidden units was 424.2 (s.d.=19.0), and receptive field sizes ranged from 4.0 to 13.4 (mean=5.68, s.d=2.2).

Synthetic Brain Imaging

Synthetic brain imaging in the models was performed by measuring for each verb the activation flowing through the direct input-output and the hidden-output pathways. It was assumed that the corresponding pathways in the brain would be spatially separated so that in brain imaging, activation differences between them could be observed. Activation in the direct input-output pathway was computed as the summed absolute activation reaching the output units through these direct connections (i.e., $\sum_o \sum_i |w_{oi} a_i|$ where o are output units, i input units, w_{oi} the weight of the connection between input unit i and output unit o, and a_i the activation of input unit i). Likewise, activation in the hidden pathway was computed as the absolute activation reaching the output units through the hidden-output connections.

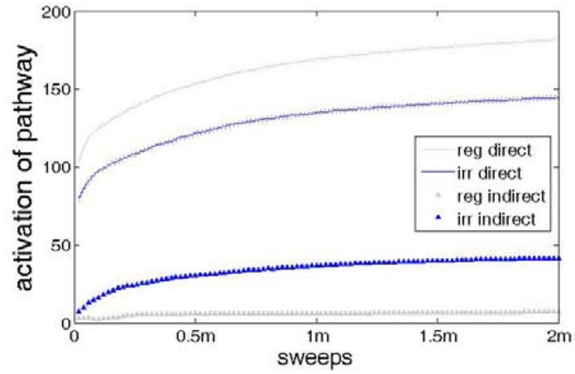

Figure 2: Development of the activation profiles of both network pathways by regular and irregular verbs. All results are averaged over 10 networks.

Figure 2 shows the developmental profile for the activation of each network pathway by regular and irregular verbs. In relation to irregular verbs regulars activated the direct pathway more, and irregulars showed more activation than regulars in the indirect pathway. This double dissociation between regular and irregular verbs was apparent early in development but became stronger as the model reached its adult state with a fully developed hidden layer and tuned connection patterns. The fact that a double dissociation in regional activation between regular and irregular verbs emerges in the model on the basis of a single associative mechanism undermines the argument put forward in defense of the dual-mechanism view that differential activation of brain regions for both verb types indicates an underlying qualitative processing difference (e.g., Jaeger et al., 1996). This result further provides evidence for the view that the source of dissociations between verbs lies in the ease of processing of past tense forms. In the model the developing hidden layer allocates processing resources for verbs that are harder to process, and the fact that this pathway is preferentially activated by irregular verbs indicates that on average, irregular forms are harder to process than regulars. The argument that dissociations between verbs reflect ease of processing has been made previously with respect to imaging studies (Seidenberg & Arnoldussen, 2003), but the present model provides a mechanistic account of how these dissociations can arise on the basis of a neuroconstructivist developmental process. Previous approaches have also based explanations for dissociations on a differential involvement of semantics in the generation of regular and irregular forms (e.g., Joanisse & Seidenberg, 1999; Patterson et al., 2001). Without precluding the possibility that semantic and irregular processing might be linked and correlated, the present model shows that it is, however, not necessary to postulate semantic involvement in the mechanism of irregular verb inflection to explain observed processing dissociations.

The model also helps to explain why different imaging studies have found very different areas being activated for regular and irregular verbs. The 'easiness' view of processing indicates that different statistical characteristics of verbs affect ease of processing independent of whether they are regular or irregular. Synthetic brain imaging offers the possibility to investigate in detail which factors affect ease of processing because the model is tested on a large set of verbs in which factors vary considerably. In order to investigate which factors affect ease we characterized each verb along a range of factors that were accessible to the model and therefore could affect learning: log past frequency, number of phonological friends, enemies, and friend/enemy ratio (types and tokens), presence of a stem final alveolar consonant, and phonological complexity (defined as number of phonemes per syllable). Two further factors that were not directly accessible to the network and whose effects therefore would arise from a combination of other factors were age of acquisition (AoA) and verb regularity.

Table 1: Correlations between statistical properties of verbs and their activation of each network pathway. Age of acquisition and regularity were not encoded in the training data. (** indicates p<.01).

Factors	Direct path	Indirect path
Friends (types)	0.032	-0.095**
Friends (tokens)	-0.024	-0.025
Enemies (types)	-0.25**	0.57**
Enemies (tokens)	-0.24**	0.41**
Friend/enemy ratio (types)	0.11**	-0.46**
Friend/enemy ratio (tokens)	0.14**	-0.37**
Past tense frequency	-0.17**	0.67**
Stem final alveolar	-0.3**	-0.16**
Phonol. Complexity	0.15**	0.1**
Age of acquisition	0.45**	-0.43**
Regularity	0.21**	-0.7**

Table 1 shows that most of the considered statistical factors correlate significantly with activation strengths in both pathways. The results indicate that low frequency verbs with an advantageous neighborhood (many friends, few enemies, more friends than enemies) tend to differentially activate the direct pathway, while the indirect pathway gets activated most for frequent verbs with unfavorable neighborhood. Phonological complexity is correlated positively with activation in both pathways, indicating that phonologically complex verbs led to overall greater activation in the model. Conversely, presence of a stem-final alveolar is correlated with lower activity across the network.

Those factors that were accessible to the network during training (all but AoA and regularity) were entered into a step-wise regression model to establish their contribution to the ratio of direct/indirect path activation (see Table 2).

Table 2: Reliable predictors of the relative activation ratio in a step-wise regression analysis.

Factors	coefficient	p-value
Past tense frequency	-0.2307	< 0.001
Enemies (types)	-0.1448	< 0.001
Friend/enemy ratio (types)	0.0732	< 0.001
Friend/enemy ratio (tokens)	-0.0636	< 0.001
Phonol. complexity	-0.0209	< 0.05
r^2 of regression model	0.64	< 0.001

These results suggest a picture that is slightly more complex than directly linking activation of the hidden pathway with low ease of processing: frequent forms as well have a higher chance of using the indirect pathway because even small errors accumulated through frequent exposure will lead to allocation of hidden units. The results also show that the effect of verb regularity that is taken as evidence for separate underlying processing mechanisms in the dual mechanism theory arises from a combination of other factors. Most regular verbs show a high ratio of direct/indirect pathway activation. The ten most typical monosyllabic regular verbs according to this measure are *tax, glaze, wax, bless, wail, draft, flinch, flare, tune,* and *wire*. The ten least typical monosyllabic regulars, that is, those regulars with the lowest direct/indirect activation ratio, are *look, need, live, seem, like, try, add, start, pick,* and *want.* Interestingly, these are verbs that would normally be regarded as prototypically regular because of their high frequencies. The synthetic imaging results presented here, however, suggest that in brain imaging studies these verbs might actually activate similar regions to irregulars.

The ten irregular verbs with the most typical irregular activation pattern, that is, a low direct/indirect pathway activation ratio, are *go, say, do, know, think, teach, see, take, seek,* and *make.* The ten least typical irregulars according to this measure are *thrust, creep, stride, freeze, stink, lend, shrink, strive, bet, and swear.*

Discussion

The model described in this paper, which has previously been used to account for developmental and adult behavioral data in past tense inflection, was here extended to model brain imaging data. The model shows how, based on a learning process in which experience-dependent brain development interacts with a structured environment, an adult processing system emerges that displays dissociations between regular and irregular verbs on the basis of a single, associative mechanism. Investigating the underlying causes for the emergence of these dissociations we have identified a number of statistical factors that affect in which region of the network a verb is preferentially processed.

Together these results raise a number of important points. First, the synthetic brain imaging results presented here – double dissociations between regulars and irregulars in regional activation patterns – have in previous brain imaging studies been taken as prime evidence for a dual-mechanism view of processing with a grammatical mental rule for regulars and associative storage for irregulars. The fact that in the model the same dissociations emerge on the basis of a single processing mechanism presents strong evidence against this account and in favor of single mechanism views of inflection processing. Second, the brain regions activated in producing the past tenses for specific verbs are likely to be a complex function of interacting statistical verb properties, with frequency of the past tense form, number of different phonological enemies, friend/enemy ratio and phonological complexity making significant contributions. Third, these results indicate that one should not expect typical regulars and typical irregulars to activate different brain regions. Instead, because both regular and irregular verbs that are regarded as typical have a high frequency, they might in fact utilize the same brain regions. With frequency in the model being one of the main factors that determine which processing resources are used, one would expect all frequent verbs to activate overlapping brain regions, and less frequent verbs showing different brain activation patterns according to their other statistical properties. It is therefore not advisable to match a small set of regular and irregular verbs on frequency and expect dissociations between them to become visible in imaging studies. Fourth, these results indicate that verbs might dissociate differently depending on experimental paradigm. Added processing resources are allocated both for frequent verbs and for verbs that are hard to process. Therefore, in imaging studies, frequent and hard verbs might activate the same brain regions. In contrast, in behavioral paradigms such as lexical decision tasks frequent and hard verbs might dissociate because interactions between processing regions can differ with specific task demands even when the same regions are involved in processing both. Computational modeling can be used to investigate these questions further.

Acknowledgments

This work was supported by ESRC grant RES-061-23-0129 to Gert Westermann.

References

Arbib, MA, Billard, A, M., I, & E., O (2000). Synthetic brain imaging: grasping, mirror neurons and imitation. *Neural Networks, 13*, 975-997.

Baayen, H, Piepenbrock, R, & van Rijn, H. (1993). The CELEX Lexical Database.

Cangelosi, A, & Parisi, D (2004). The processing of verbs and nouns in neural networks: Insights from synthetic brain imaging. *Brain and Language, 89*, 401-408.

Desai, R, Conant, LL, Waldron, E, & Binder, JR (2006). FMRI of past tense processing: The effects of phonological complexity and task difficulty. *Journal of Cognitive Neuroscience, 18*, 278-297.

Dhond, RP, Marinkovic, K, Dale, AM, Witzel, T, & Halgren, E (2003). Spatiotemporal maps of past-tense verb inflection. *Neuroimage, 19*, 91-100.

Jaeger, JJ, Lockwood, AH, Kemmerer, DL, VanValin, RD, Murphy, BW, & Khalak, HG (1996). Positron Emission Tomographic Study of Regular and Irregular Verb Morphology in English. *Language, 72*, 451-497.

Joanisse, MF, & Seidenberg, MS (1999). Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences of the United States of America, 96*, 7592-7597.

Joanisse, MF, & Seidenberg, MS (2005). Imaging the past: Neural activation in frontal and temporal regions during regular and irregular past-tense processing. *Cognitive Affective & Behavioral Neuroscience, 5*, 282-296.

Just, MA, Carpenter, PA, & Varma, S (1999). Computational modeling of high-level cognition and brain function. *Human Brain Mapping, 8*, 128-136.

Lavric, A, Pizzagalli, D, Forstmeier, S, & Rippon, G (2001). A double-dissociation of English past-tense production revealed by event-related potentials and low-resolution electromagnetic tomography (LORETA). *Clinical Neurophysiology, 112*, 1833-1849.

Li, P, & MacWhinney, B (2002). PatPho: A phonological pattern generator for neural networks. *Behavior Research Methods, Instruments, & Computers, 34*, 408-415.

Mareschal, D, Johnson, MH, Sirois, S, Spratling, MW, Thomas, M, & Westermann, G (2007). *Neuroconstructivism: How the Brain Constructs Cognition*. Oxford, UK: Oxford University Press.

McClelland, JL, & Patterson, K (2002). Rules or connections in past-tense inflections: what does the evidence rule out? *Trends in Cognitive Sciences, 6*, 465-472.

Patterson, K, Lambon Ralph, MA, Hodges, JR, & McClelland, JL (2001). Deficits in irregular past-tense verb morphology associated with degraded semantic knowledge. *Neuropsychologia, 39*, 709-724.

Pinker, S (1999). *Words and Rules: The Ingredients of Language*. New York: Basic Books.

Pinker, S, & Ullman, MT (2002). The past and future of the past tense. *Trends in Cognitive Sciences, 6*, 456-463.

Ruh, N, & Westermann, G (2008). *A single-mechanism dual-route model of German verb inflection.* Paper presented at the Proceedings of the 30th Annual Conference of the Cognitive Science Society.

Seidenberg, MS, & Arnoldussen, A (2003). The brain makes a distinction between hard and easy stimuli: Comments on Beretta et al. *Brain and Language, 85*, 527-530.

Seidenberg, MS, & Hoeffner, JH (1998). Evaluating Behavioral and Neuroimaging Data on Past Tense Processing. *Language, 74*, 104-122.

Ullman, MT, Corkin, S, Coppola, M, Hickok, G, Growdon, JH, Koroshetz, WJ, & Pinker, S (1997). A neural dissociation within language: Evidence that the mental dictionary is part of declarative memory, and that grammatical rules are processed by the procedural system. *Journal of Cognitive Neuroscience, 9*, 266-276.

Westermann, G (1998). Emergent modularity and U-shaped learning in a constructivist neural network learning the English past tense. In M. A. Gernsbacher & S. J. Derry (Eds.), *Proceedings of the Twentieth Annual Conference of the Cognitive Science Society*. Hillsdale, NJ: Erlbaum.

Westermann, G (2000). A constructivist dual-representation model of verb inflection. In L. R. Gleitman & A. K. Joshi (Eds.), *Proceedings of the Twenty-Second Annual Conference of the Cognitive Science Society*.

Westermann, G, Mareschal, D, Johnson, MH, Sirois, S, Spratling, M, & Thomas, M (2007). Neuroconstructivism. *Developmental Science, 10*, 75-83.