

The evolution of language: A neurolinguistic perspective¹

Marco Haverkort
University of Nijmegen & Boston University

&

Laurie A. Stowe
University of Groningen

1. Introduction

A number of different sources of evidence have been brought forward in speculating on the evolutionary roots of the human language ability, ranging from typological and historical linguistics to archeological observations and the study of ritual and poetry. The different contributions to this volume offer a representative cross-section. In this paper, we want to explore to what extent the organization of language in the brain² sheds light on the question of its evolutionary origin.

I will argue that the areas that have traditionally been associated with language, especially Broca's area, are not language-specific and that areas that are primarily important in supporting cognitive functions other than language turn out to play a crucial role in language processing, too. These observations support the idea that language is the result of a process of EXAPTATION, rather than adaptation (as maintained by Pinker & Bloom 1990 and Pinker 1995, for instance).

Modern neuro-imaging techniques allow the study of the brain in vivo in normal, unimpaired subjects, thus excluding 'rewiring' as a consequence of brain damage. These techniques have a number of additional advantages over the study of patient populations. They are noninvasive, and they allow studying areas and functions, such as language, that are rarely selectively impaired in humans or for which animal models are not informative. Moreover, they allow the study of temporal and spatial dynamics of neural activity. Lesions are rarely well circumscribed; the spatial resolution of a number of neuro-imaging methods allows for a more precise localization of distinct functions (Haverkort 2001).

2. Positron Emission Tomography

The studies reported in this paper utilize Positron Emission Tomography (PET). PET is a technique where subjects are injected with radioactively labeled water (H_2^{15}O).³ Metabolic changes occur in the parts of the brain that are activated

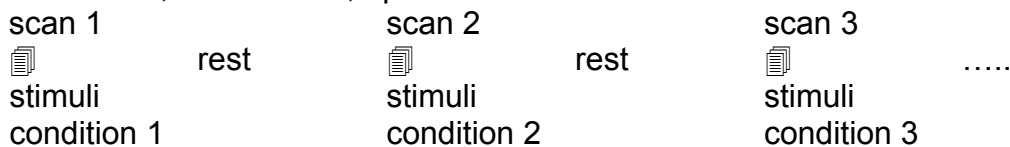
¹ The research of Dr. Haverkort has been made possible by a research fellowship of the Royal Netherlands Academy of Arts and Sciences.

² For a more in-depth discussion of some of the issues presented here, the reader is referred to Stowe et al. (2001).

³ Alternatively, individual neurotransmitters can be labeled with a radioactive tracer, providing detailed information about specialized functions of these neurotransmitters in specific areas of the brain.

during a task, causing changes in regional cerebral blood flow patterns, the so-called hemodynamic responses. In areas where the blood supply increases, the amount of gamma radiation increases as well as a function of the changes in blood flow. This technique provides the researcher with a three-dimensional representation of the brain.

As this technique measures hemodynamic changes, it provides only an indirect measure of neural activity. Hemodynamic responses have the disadvantage that they are relatively slow: It takes about 6 seconds to reach a peak in blood flow in an activated area after the onset of activation. Thus, the temporal resolution of PET is rather low⁴ and stimuli have to be presented for longer periods of time, in order for hemodynamic changes to be reliably measured. As a consequence of this, the design used in PET studies is a so-called 'blocked design', where stimuli for a condition are presented over a period of about 40 seconds, followed by a rest condition of about 15 minutes (necessary to allow the radiation level to drop to background level⁵), followed by stimuli for another condition, and so forth, up to a maximum of 12 conditions/scans.



Stimuli are presented visually – one word at a time at the center of a computer screen, to avoid noise associated with eye and head movements. Subjects participating in the experiments presented here were all right-handed and had normal or corrected to normal vision.

3. Broca's area: Syntax and language production

Ever since Broca's (1861) seminal paper, the inferior frontal gyrus (Brodmann's area 44) has been associated with language. It has been assumed to play a role in language production, and more specifically, be crucial for syntax, as patients with lesions in this area have problems producing language, while their understanding seems pretty much intact.⁶ Both of these ideas are misguided.

Stowe et al. (2001) show that Broca's area is also activated during language comprehension. Subjects were asked to read sentences and word lists. It turned out that Broca's area was not reliably activated by simple, one-clausal sentences, but it was activated by more complex sentences which contained (center and right) embeddings or syntactic ambiguities. Surprisingly, this area was also activated when subjects read word lists. These results indicate that Broca's area is not solely activated in language production, but also in language

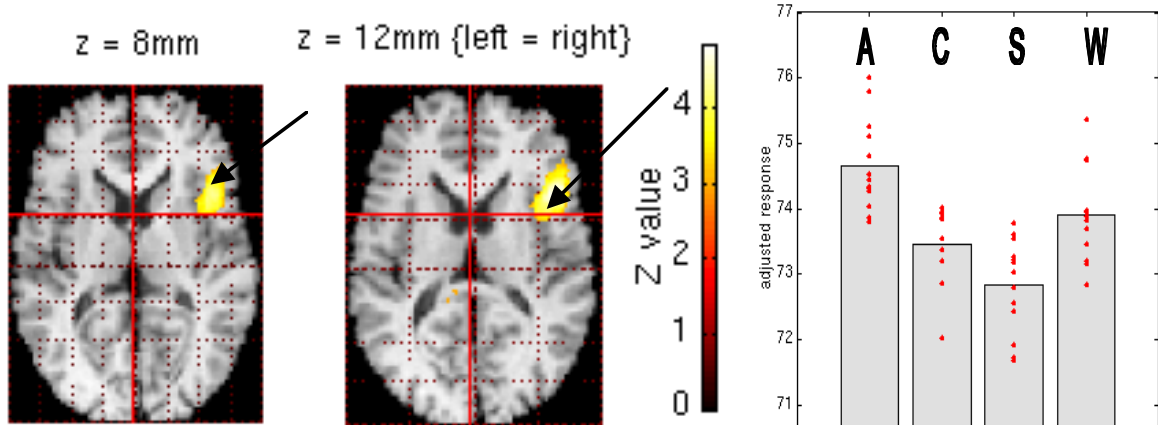
⁴ Spatial resolution of PET makes up for the poor temporal resolution. Deviations are in the order of magnitude of the millimeter.

⁵ Half-life of the tracers used is about 2 minutes.

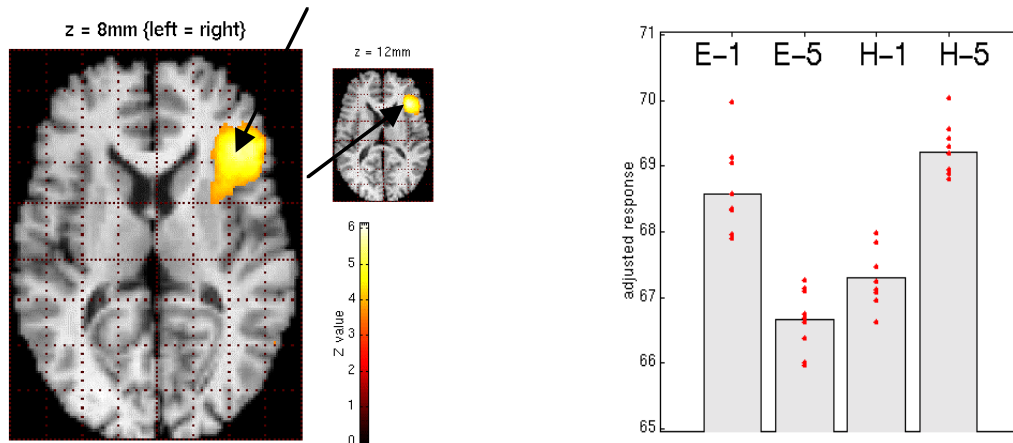
⁶ Under experimental conditions, though, these patients have problems understanding semantically reversible passives, object clefts, and object relatives, while their understanding of the parallel actives, subject clefts, and subject relatives is relatively unimpaired (Caplan & Hildebrandt 1988, Grodzinsky 1990). This indicates that Broca's area plays a role in language comprehension.

comprehension, and that it is activated in conditions that do not involve any syntax, viz. word lists.

The activations are shown in the left panel of the following figure, projected onto anatomical MRI slices to give an indication of their anatomical location. Each slice is viewed as if from below with the anterior part of the brain upwards and the posterior part downwards, so that the left hemisphere appears on the right side of the image and vice versa. In the right panel, the relative regional blood flow (in milliliters of blood per minute per deciliter brain volume) is shown for the four conditions included in the comparison: A= Ambiguous sentences, C=Complex sentences, S=Simple sentences, and W=Word lists.



The amount of activation of this area can be characterized in terms of a combination of lexical and phrasal working memory load (both high in complex and ambiguous sentences, relative to simple sentences): The more these increase, the more the area is activated. One could still maintain that the activation related to working memory forms a distinct neural network from the network involved in syntactic processing, which both happen to be located in the same general area, conform the traditional view. This interpretation of the data can easily be discarded, though. If a different neural network were involved in memory tasks and in syntactic processing, there should be no interaction between the two tasks, as they would not compete for the same (limited) resources. To investigate this, subjects were presented with simple and complex sentences and simultaneously had to monitor these sentences for one or five words that they were given before the experiment started. Thus, there were 4 combinations: simple sentences (E) with a memory load of one word and a memory load of five words, and complex sentences (H) with the same respective memory loads. Again, the left inferior frontal gyrus was activated in these tasks, as shown in the following figure.



The area in the left inferior frontal gyrus in which syntactic complexity and memory load interact is shown in the left panel (note: the left hemisphere appears on the right and vice versa). The amount of activation for each condition is shown in the right panel (E-1=Easy Sentences/Low Load; E-5=Easy/High; H-1=Hard/Low; H-5=Hard/High). Interestingly, syntactic complexity and memory load interact; there is no simple additive effect. This clearly indicates that the same system underlies both tasks, thus refuting the idea that syntactic processing and memory tasks make use of distinct neural networks. Had that been the case, then the two tasks would have had an additive effect on regional cerebral blood flow. Positive correlations between speed and accuracy of syntactic processing on one hand and memory span on the other have been reported in the behavioral literature, independently supporting the idea that the same resources underlie these two tasks (Daneman & Carpenter 1980, Just & Carpenter 1992).

Importantly, Broca's area cannot be regarded as specifically and solely supporting linguistic processing. Recent studies have shown that the perception of rhythm or imagining the completion of a tune also activate the left inferior frontal gyrus (for instance, Platel et al. 1997 and Halpern & Zatorre 1999). Language and music share the incremental hierarchical structure of their syntax (Lerdahl & Jackendoff 1983, Staal 1989). Thus, the left inferior frontal gyrus seems to be involved in processing this type of structure, independently of its specific nature – linguistic or musical – or of its modality – motor or auditory.

4. Additional activations: Visual cortex, cerebellum, and right fronto-temporal region

In the experiment just described, additional activation in the visual cortex could be observed, which we interpret as a result of an overflow in case the working memory capacity of Broca's area is exceeded. Independently, it could be shown that the visual cortex is activated as a function of syntactic complexity by comparing the hemodynamics of subject and object relative clauses in Dutch:

- (1) Subject relative:

Anna misleidde de psycholoog; die t_i de spelers ophemelde
 Anna deceived the psychologist who the players praised-3S
 'Anna deceived the psychologist who praised the players'

(2) Object relative:

Anna misleidde de psycholoog; die de spelers t_i ophemelden

Anna deceived the psychologist who the players praised-3P

'Anna deceived the psychologist whom the players praised'

As these examples show, subject and object relative clauses in Dutch, unlike their English counterparts, have the same word order; they are only disambiguated by the inflectional ending on the sentence-final verb (-*e/-en*). Object relatives are harder to process than subject relatives: Subjects take longer reading them in self-paced reading tasks, they make more regressive eye movements while reading them, and they make more mistakes in interpreting them (Mak 2001). This additional processing load is reflected by a hemodynamic response in the visual cortex in the occipital lobe.⁷ As with the task simultaneously varying syntactic complexity and memory load, this effect can be interpreted as an overflow due to a working memory load that exceeds the capacity of Broca's area.

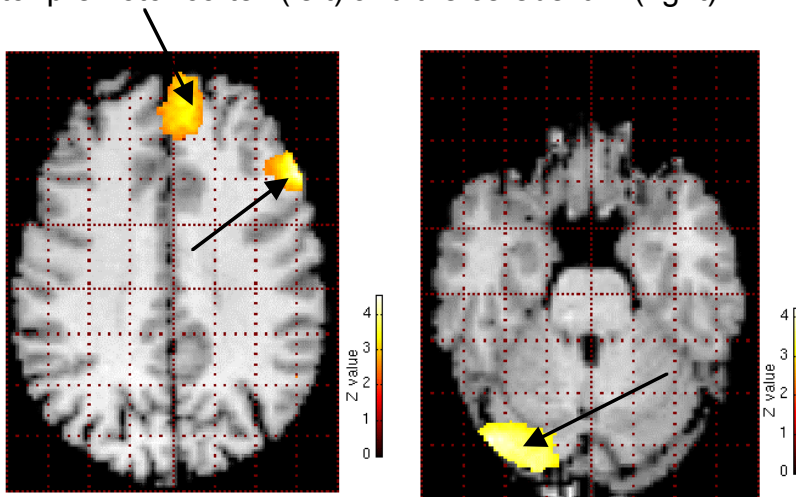
Areas found to be activated by syntactically ambiguous sentences relative to unambiguous control sentences involve the right cerebellum and the left motor cortex, which are both primarily associated with motor control.

Zij kunnen bakken met zulk deeg niet verplaatsen

they can bake(V)/containers(N) with such dough not move

'They can't move containers with such dough'

In this sentence, the unmarked syntactic analysis is the one where *bakken* is interpreted as an infinitival verb form governed by the modal auxiliary. It is only when the negation is encountered – which in Dutch immediately follows the finite verb in main clauses – that it becomes clear that this analysis is impossible and that *bakken* needs to be interpreted as a noun, causing a syntactic reanalysis. The following figure shows the associated activation for the motor/premotor cortex (left) and the cerebellum (right):



⁷ This effect cannot be a reflection of the process of reading the sentences, as both sentence types were presented in the same way. Hence the hemodynamic patterns associated with the processing underlying the reading of the two sentence types cancel each other out. The observed effect can only reflect the additional processing load associated with the object relatives.

Co-activation of the right cerebellum and the left motor cortex is neuroanatomically plausible, as the right cerebellum is connected to the left motor cortex. However, no facial movement was involved, as the contralateral homologues of these two areas were not activated. In the literature, these areas have been argued to involve an articulatory rehearsal mechanism (Baddeley 1986, Gathercole & Baddeley 1993), but the activation found in this experiment is only unilateral.

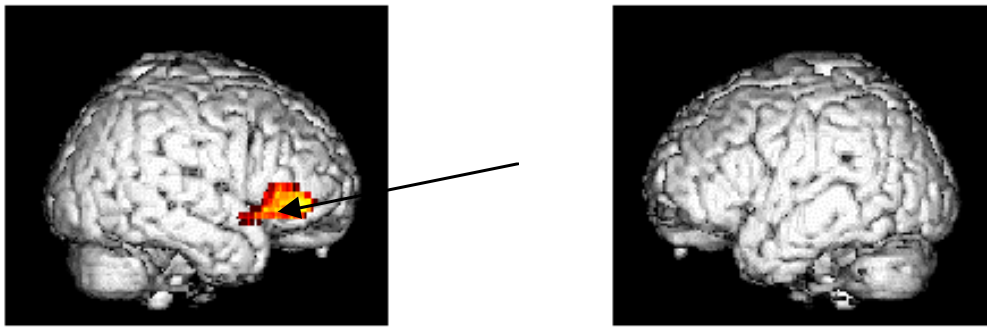
Fiez et al. (1992), reporting on the performance of a patient with a right cerebellar lesion in a verbal fluency task, found the patient to be prone to making mistakes involving the production of highly associated words that did not fit the criteria of the task. The patient's performance did not improve to a semiautomatic, nearly perfect performance level with experience in the task, as it did in control subjects. Thus, the cerebellum plays a crucial role in error detection. An error detection mechanism is indispensable in parsing garden path sentences, where syntactically ambiguous sentences are resolved to a less-preferred reading, thus forcing a revision of the initial analysis.

The left hemisphere has traditionally been assumed to be the locus of language processing, but evidence amasses that the right hemisphere contributes to the process as well. The right fronto-temporal region is activated in the process of lexical ambiguity resolution.⁸ Subjects were given sentences like the following:

- (4) De ezel stond in de schuur al lang te rotten
the donkey/easel stood in the shed already long to rot

After the occurrence of a lexically ambiguous word, *ezel* in the above example, the sentence remained neutral as to which meaning of this word was the correct one. Ultimately, however, the sentence context forced the preferred or the non-preferred meaning ('easel', in this case). When the nonpreferred meaning turned out to be the correct one, subjects were forced to abandon the preferred meaning they had been assuming and activate and integrate the alternative meaning which they had been actively inhibiting until then; when the preferred meaning turned out to be the correct one, subjects did not need to maintain the nonpreferred alternative any longer. The right fronto-temporal lobe appears to support maintenance of a secondary meaning rather than revision of meaning – in contrast to syntactic ambiguity, in which structural revision occurs (see above) –, as it is activated when the sentence favors both the nonpreferred and the preferred meaning, i.e. when revision would be forced and when it would not be necessary, respectively. If revision were involved, activation of the fronto-temporal region should only take place in the former case, when the sentence was resolved towards the nonpreferred meaning, contrary to fact (Haverkort & Stowe 2001).

⁸ There is also evidence from patient studies supporting this. Patients with frontal right hemisphere damage have a hard time producing the second, nondominant meaning of an ambiguous word. Additionally, right hemisphere damage has been associated with the incapacity to understand any form of nonliteral language (idioms, metaphor), as well as an inability to maintain adequate discourse representations and make any inferences.



This activation was more pronounced for ambiguous verbs than for ambiguous nouns, an effect that can be explained by the more complex nature of the lexical representation of verbs, involving syntactic information such as argument structure. Crucially, though, as the above figure shows, no area in the left hemisphere (on the right) was activated in this task, supporting the conjecture that only maintenance of a secondary meaning is involved.

In sum, language processing involves a large number of specialized areas, forming a complex distributed network. All the areas involved are not primarily or not solely associated with linguistic processing. Broca's area, a traditional language area, is not solely activated in sentence production or when syntactic processing is involved; moreover, it is activated in other than linguistic tasks that involve temporary storage of incremental hierarchical information, as when processing musical structure. The cerebellum, motor and visual cortices, which are primarily responsible for planning and coordinating, executing complex movements, and processing of visual stimuli, respectively, become involved in language processing as the syntactic complexity (hence working memory load) of the stimuli increases. All the areas discussed above thus have a primary function which is nonlinguistic in nature or a function that underlies linguistic as well as other cognitive abilities. This observation will be argued to be crucial in the debate between adaptation and exaptation.

5. Adaptation vs. exaptation

One of the important debates in evolutionary biology focuses on the question whether traits in a population should be explained in terms of adaptation or exaptation. Adaptation holds that random mutations are directly favored by natural or sexual selection in a particular context. Natural selection favors traits that enhance the survival value of an individual; individuals who lack that trait, or have it to a lesser extent, have a smaller chance of surviving in their natural environment. The population will hence show an increase in a specific trait because the members that possess the trait in the most pronounced manner stand a better chance of surviving and reproducing, thus passing on the relevant genes to their offspring.

In a similar manner sexual selection may contribute to an increase of a specific trait in a population. But this time it is not because it increases survival chances, but because members of the opposite sex are more attracted to individuals that possess that trait than to individuals that do not, hence enhancing

mating chances of the former. Sexual selection favors traits that signal health and strength but that are not necessarily important in terms of natural selection; they can sometimes even endanger survival of an individual. An example would be the male peacock's tail. A long, colorful tail attracts the hens by signaling health, but it decreases the male's chances of survival: Because it is so colorful, it also attracts the attention of predators, and because it is so long, it will make the male less agile and hence an easier prey for predators.

The adaptation/exaptation debate in evolutionary biology focuses on the question whether traits are directly selected for or whether selection is indirect. Adaptation is a form of direct selection: A phenotypical trait enhances survival chances or mating chances of an individual and is therefore selected. Thus an adaptation is 'the evolutionary modification of a character under selection for efficient or advantageous (fitness-enhancing) functioning in a particular context or set of contexts' (West-Eberhard 1992: 13). Exaptation, on the other hand, is an indirect form of selection: A phenotypical trait is selected for a specific function, but then turns out to be useful for a completely different function and ends up being co-opted for that new function, thereby often losing its original function (Gould & Lewontin 1979, Gould & Vrba 1982).

The evolution of language, more in particular of syntax, is very different under these two viewpoints. Adaptationism maintains that communication presented a selectional advantage – in terms of natural or sexual selection (cf. Miller (2001) for the latter) – and that therefore individuals who were good communicators had a better chance of surviving and passing on their genes than individuals who did not have this property, or had it only to a lesser extent. Gradually, the communication system became formally more and more complex, resulting in an increase in functional complexity, i.e. what could be expressed with it, by introducing, for instance, a tense, mood, and aspect system, so that communication was not limited to the here and now.

Under an exaptationist perspective, on the other hand, syntax developed for a different reason than communication and only later turned out to be useful for communication (see also Staal 1989, de Roder (this volume)). One logical argument one can bring to bear on this issue is that language is not an optimal means of communication: Ambiguities at the lexical and syntactic level abound and in general there is no isomorphism between form and meaning. Not surprisingly, humans have to take recourse to formal logic if they want to be completely unambiguous and maintain a form-meaning isomorphism. If language, more in particular syntax, had developed because of communication, one would expect a development towards less ambiguity, contrary to what can be observed.

The neurolinguistic evidence discussed above suggests that the system has developed in a nonadaptive, exaptationist fashion, integrating different neural subsystems that could be profitably co-opted for a new function. These different subsystems – involving a number of different aspects of storage and computation – turned out to be useful for language processing even though most of them were not primarily or solely adapted to perform this specific task: Broca's area is not only activated by language but also by music; it is activated in explicit memory tasks, as well as language processing; the right hemisphere, which was believed to lack linguistic capacities up until recently, turns out to

play a crucial role; and areas that are primarily related to cognitive functions other than language, such as the visual cortex, the motor cortex, or the cerebellum, play a role as well. In their new combination, complexity was able to emerge, increase, and stabilize (cf. Casti 1994, Holland 1995, Lorenz 1982). Interestingly, the different subsystems involved did not lose their original functions, as often happens in the process of exaptation; they were only partially exapted, contributing to an additional process while still preserving their original function, which was so central to survival.

The knowledge base, in the form of a grammar, and the processing mechanisms that use that grammar should be clearly distinguished,⁹ although they crucially interact in development. The processing mechanisms impose boundary conditions on the complexity of the developing grammatical knowledge. The complexity of grammars cannot exceed the processing capacity of the brain, as that would make them useless (Haverkort & Neijt 2001). Thus, qualitative changes in the grammars, such as the emergence of recursion, can be assumed to have only arisen when the relevant boundary conditions – in terms of working memory capacity or processing speed, for instance – were fulfilled.

A consequence of the view presented here, is that language is not modular, in the sense that the apparatus supporting language processing is not domain-specific or informationally encapsulated (contra Fodor 1983).¹⁰ Most areas involved in language processing are associated primarily or partially with other cognitive functions, such as vision or motor planning. Modules, under this view, obtain a much more abstract status under evolutionary pressure, being at least to some extent less modality- and function-specific than has hitherto been assumed.

References

- Baddeley, A. 1986. *Working memory*. Oxford: Oxford University Press.
- Broca, P. 1861. Remarques sur le siège de la faculté du langage articulé, suivies d'une observation d'aphémie (perte de la parole). *Bulletin de la Société Anatomique* 6.
- Caplan, D. & N. Hildebrandt. 1988. *Disorders of syntactic comprehension*. Cambridge: MIT Press.
- Casti, J. 1994. *Complexification: Explaining a paradoxical world through the science of surprise*. New York: Harper Collins.

⁹ The behavior of aphasics provides evidence supporting such a distinction. Patients' behavior fluctuates from day to day and there is often considerable post-onset improvement without re-acquisition. Moreover, patients exhibit different behavior depending on the task demands (sentence-picture matching vs. grammaticality judgments, for instance), while the syntactic constructions under investigation remain the same. Finally, syntactic priming effects can be shown to exist when the stimulus-onset asynchrony is increased, indicating again that the relevant knowledge is available under certain conditions. The patients seem to have an online processing problem; there is a mismatch between the complexity of their syntactic knowledge on one hand, and the limited capacity of their language processing mechanisms and working memory on the other, precisely as discussed here for evolution (cf. also Haverkort 1993).

¹⁰ The heavy reliance on notions from folk psychology in Fodor (1983) probably explains why these were introduced as defining characteristics of modules.

- Daneman, M. & P. Carpenter. 1980. Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*.
- de Roder, J. this volume. Poetry: The missing link.
- Fiez, J., S. Petersen, M. Cheney & M. Raichle. 1992. Impaired non-motor learning and error detection associated with cerebellar damage. *Brain* 115.155-78.
- Fodor, J. 1983. *The modularity of mind*. Cambridge: MIT Press.
- Gathercole, S. & A. Baddeley. 1993. *Working memory and language*. Sussex: Psychology Press.
- Gould, S. & R. Lewontin. 1979. The spandrels of San Marco and the Panglossian paradigm. *Proceedings of the Royal Society of London*, B 205.
- Gould, S. & E. Vrba. 1982. Exaptation: A missing term in the science of form. *Paleobiology* 8.
- Grodzinsky, Y. 1990. *Theoretical perspectives on language deficits*. Cambridge: MIT Press.
- Halpern, A. & R. Zatorre. 1999. When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex* 9.
- Haverkort, M. 1993. Universalgrammatik, Parameter und Agrammatismus. *Linguistische Berichte* 143.
- Haverkort, M. 2001. *Neuroimaging methods*. Paper presented at the 11th Conference of the European Second Language Association, University of Paderborn.
- Haverkort, M. & A. Neijt. 2001. *Complexiteit in taal en evolutie*. University of Nijmegen, ms.
- Haverkort, M. & L. Stowe. 2001. *Neural mechanisms of lexical ambiguity resolution*. Paper presented at the Neurological Basis of Language Conference, University of Groningen.
- Holland, J. 1995. *Hidden order: How adaptation builds complexity*. Reading, MA: Addison-Wesley.
- Just, M. & P. Carpenter. 1992. A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*.
- Lerdahl, F. & R. Jackendoff. 1983. *A generative theory of tonal music*. Cambridge: MIT Press.
- Lorenz, K. 1982. *Vergleichende Verhaltensforschung: Grundlagen der Ethologie*. Munich: DTV Wissenschaft.
- Mak, P. 2001. *Processing relative clauses: Effects of pragmatic, semantic and syntactic variables*. Nijmegen: University of Nijmegen dissertation.
- Miller, G. 2001. *The mating mind: How sexual choice shaped the evolution of human nature*. London: Random.
- Pinker, S. 1995. Facts about human language relevant to its evolution. *Origins of the human brain*, ed. by J.-P. Changeux & J. Chavaillon. Oxford: Clarendon.
- Pinker, S. & P. Bloom. 1990. Natural language and natural selection. *Behavioral and Brain Sciences* 13.
- Platel, H., C. Price, J.-C. Baron, R. Wise, J. Lambert, R. Frackowiak, B. Lechevalier & F. Eustache. 1997. The structural components of music perception: A functional anatomical study. *Brain* 120.
- Staal, F. 1989. *Rules without meaning: Ritual mantras and the human sciences*. New York: Peter Lang.

Stowe, L., M. Haverkort & F. Zwarts. 2001. *Rethinking the neurological basis of language*. University of Groningen, University of Nijmegen, and Boston University, ms.

West-Eberhard, M. 1992. Adaptation: Current usages. *Keywords in evolutionary biology*, ed. by E. Keller & E. Lloyd. Cambridge: Harvard University Press.