



HAL
open science

Review article: The imaging of what in the multilingual mind?

Kees de Bot

► **To cite this version:**

Kees de Bot. Review article: The imaging of what in the multilingual mind?. *Second Language Research*, 2008, 24 (1), pp.111-133. 10.1177/0267658307083034 . hal-00570740

HAL Id: hal-00570740

<https://hal.science/hal-00570740>

Submitted on 1 Mar 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Review article

The imaging of what in the multilingual mind?

Kees de Bot *University of Groningen*

Received December 2006; revised February 2007; accepted July 2007

In this review article it is argued that while the number of neuro-imaging (NI) studies on multilingual processing has exploded over the last few years, the contribution of such studies to enhance our understanding of the process of multilingual processing has not been very substantial. There are problems on various levels, which include the following issues: ownership of the field of NI and multilingualism, whether relevant background characteristics are assessed adequately, whether we consider variation as a problem or a source of information, what NI tells us about multilingual development, whether the native speaker is the norm in NI research, the added value of NI data, what information NI provides, and what the contribution of NI research is to theories about the multilingual brain. The conclusion is that as yet NI has not fulfilled the high expectations raised by the technical progress and the large number of studies that have been carried out.

Keywords: Neuro-imaging, multilingualism, ERP, research methodology, psycholinguistics, dynamic systems theory

Gullberg, M. and Indefrey, P. 2006: *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell Publishing. 348 pages. US\$ 40. ISBN 978-1405155427

I Introduction

The title of this contribution is a somewhat veiled reference to a review by Wolfgang Klein on the role of introspection in the study of second language acquisition (Klein, 1989). In this review he argues that it is

Address for correspondence: Kees de Bot, Department of Applied Linguistics, Faculty of Arts, University of Groningen, P.O. Box 716, 9700 AS Groningen, The Netherlands; email: C.L.J.de.Bot@rug.nl

very unclear what it is that we tap into when using introspective methods. Along similar lines I want to argue that the wealth of research on neuro-imaging and multilingualism so far has contributed little to a further understanding of processes of multilingual processing and second language development and indeed, that it is as yet very unclear what it is that various neuro-imaging (NI) techniques tell us about language processing in multilinguals. The framework used to put my arguments forward is a review of issues discussed in a volume on the cognitive neuro-science of second language acquisition, edited by Marianne Gullberg and Peter Indefrey (2006). This volume is based on the contributions to a symposium in 2005 in Nijmegen, the Netherlands. In his opening remarks for this symposium, Wolfgang Klein voiced concerns about the potential of cognitive neuroscience and, in particular, the available neuro-imaging techniques to explain the complex process of second language acquisition, which seems to be an extension of the doubts he raised with respect to introspection in 1989.

This is not a normal review article on a book in that it is not meant to be an evaluation to help potential buyers. For one thing, that would imply self-evaluation since the author of this review has also contributed to the book. More importantly, the issues dealt with are broader than the ones presented in this book. This book provides a wonderful opportunity to get a full picture of what is going on in this very active and dynamic field. So I do recommend this book, which will be one of the milestones in this area for a long time. In their introduction, the editors of the volume claim that the contributions in the book represent the state of the art in research on NI and multilingualism. So, this book allows us to take stock of the wealth of research that has been carried out over the last decade, and to see where we are and where we are going.

There can be little doubt that the area of neuro-imaging and second language acquisition is flourishing: a Google Scholar search using neuro-imaging and second language acquisition as search terms yielded more than 7000 hits. Bilingualism and second language acquisition feature prominently in even the highest ranking journals in the field. Given the effort put into it and the eagerness of groups of researchers who had not done any research in this field up to that point to jump on the bandwagon, one would expect this to have boosted our understanding of how multiple languages are processed in the brain. My main point is

that this is not yet the case. It could be argued that the study of multilingualism using NI is still young and that it is therefore too early for an evaluation. But on the basis of the present state of affairs it is not obvious that the situation will improve soon.

The scepticism expressed in this article is neither new, nor original. In a recent issue of *Cortex* some leading cognitive psychologists (Coltheart, 2006; Page, 2006) have expressed serious doubts about the gains made in the last decade through the application of NI techniques, and their concerns are shared by several other researchers (Van Orden and Paap, 1997; Uttal, 2001; Harley, 2004). Page's main argument is that cognitive models cannot be tested using NI data because the functional role of parts of the brain is as yet not sufficiently understood, and that NI data provide no evidence to differentiate between possible models:

By the time that cognitive models are sufficiently well specified to be able to make genuinely necessary and differential predictions regarding, say, the blood-oxygen-level dependent (BOLD) signal, then they will very likely already have enough behavioral clout to be distinguished without reference to neuroimages. (Page, 2006: 2)

In this review I point to a number of issues that I consider to be problematic. The issues discussed are the following:

- ownership of the field of NI and multilingualism;
- whether relevant background characteristics are assessed adequately;
- whether we consider variation as a problem or a source of information;
- what NI tells us about multilingual development;
- whether the native speaker is the norm in NI research;
- the added value of NI data, what information NI provides; and
- what the contribution of NI research is to theories about the multilingual brain.

II Who owns the field: inductive vs. deductive approaches

As Osterhout *et al.* (2006) indicate, in the early phases of research in multilingualism and NI, researchers seemed to be preoccupied with finding out the neural substrate for different languages, but the findings on this are still unclear. The next phase of this research focused on online processes of language production and perception. In this new phase we see an old debate in the sciences re-emerge: the contrast between an inductive approach ('research then theory') and a deductive approach

(‘theory then research’). In the Gullberg and Indefrey book, this debate comes out most clearly in the Rodriguez-Fornells *et al.* (2006) contribution and the commentary on this article by Dijkstra and van Heuven (2006). The discussion in those articles is about the presence or lack of an overall theoretical model for the large set of data provided by Rodriguez-Fornells *et al.* The deeper discussion concerns the fundamental question of whether we should have a theory of language processing (and accordingly of multilingual processing) and then test the hypotheses generated by such a theory, or whether we should rather collect the data first and then build a theory on that by having our experimental participants do various language related tasks. Using these tasks it can be shown what activity can be detected in different parts of the brain and when that activity takes place. The next question then is who is responsible for the provision of the theory to be tested. Dijkstra and van Heuven refer to a statement by Grosjean *et al.* (2003) that the collaboration between cognitive neuroscientists and psycholinguists ‘should not be a one-way street with neuroscientists proving theories devised by language scientists’ (161–62). Dijkstra and van Heuven’s reaction is ‘Why would neuroscientists reinvent the cogs of cognitive science?’

There are a number of noteworthy aspects in this discussion. A minor one is that Dijkstra and van Heuven appear to take the position that cognitive scientists are the language scientists. Obviously many linguists would take offence at such a claim. In particular the nativist school would claim that they have more to offer than just adequate linguistic descriptions (Koopmans, 2006). There is clearly more to language and language use than what has been covered so far with experimental research of the cognitive science type. But that is not the main issue. The core issue is whether NI research on multilingual processing should be based on an inductive or a deductive approach. The time of a simple specialization of language areas in the brain with clear demarcations of what Broca’s and Wernicke’s areas are supposed to do is clearly over.

The inductive approach has at any rate revealed that many more parts of the brain are involved in language processing than previously assumed. A good example is the Stowe *et al.* (2005) article, which showed that several parts of the brain that so far had not been shown to be relevant to language processing may play a role. The hedge ‘may play a role’ is used here intentionally because the fact that there is activity in a part of the brain as

evidenced by NI techniques based on metabolic activity does not necessarily imply that these areas are also relevant or needed for that activity. This may be a sign of the degeneracy referred to by Green *et al.* (2006) in their contribution: different processing mechanisms may lead to the same result without there being a necessary link. Or they may reflect activity within a brain area that does not contribute to the processing of information in other areas, as argued by Page (2006: 3). The main problem for the deductive approach is that such an approach assumes a fairly detailed understanding of the functions and roles of different parts of the brain, which we simply do not have at the moment. Dijkstra and van Heuven (2006: 195) seem to fall into this trap when they argue that '[w]ithin the BIA + framework [the bilingual processing model the authors developed], predictions can be formulated about the more specific functions of certain brain areas'. One would thus expect a list of brain areas and their functions, but such a list is not provided. What they probably mean is that the functional distinctions in their model should have some neural substrate and that there should therefore be areas that perform those functions.

By taking this position, Dijkstra and van Heuven seem to line up with what Hagoort (2006: 93) refers to as the classical view among neuro-anatomists which is that 'architectural differences in the brain structure are indicative of functional differences and, conversely, that functional differences demand differences in architecture'. Hagoort concludes from his reading of the literature that a different approach is needed. Brain areas seem to be less specialized than was assumed in the classical approach. He argues that specialization follows from use and input:

Functional differences between brain areas are in this perspective mainly due to variability of the input signals in forming functional specializations. Domain specificity of a particular piece of cortex might thus not so much be determined by heterogeneity of brain tissue, but by the way in which its functional characteristics are shaped by the input. (Hagoort, 2006: 94)

So, as things stand, the inductive approach seems to be preferred over the deductive approach. The integrative framework Dijkstra and van Heuven call for does not seem to be available or, when it is available at the functional level, the correlation with neural substrates is problematic.

That this research goes through cycles on induction and deduction probably signals the developmental stage it is in. As Indefrey (personal communication) argues: 'as soon as the functional role of at least one brain area has been identified with some certainty, deductive work can

start.’ But then the problem of who decides on that certainty remains. It is not yet clear what the boundaries are and what new types of data will become available, nor is it clear to what extent existing theories will have enough content to accommodate the data that NI research is providing.

III The localization of languages in the brain: the new Loch Ness Monster?

From the beginning of the NI research on multilingualism the localization of languages has been a major issue. In a way, the quest for the locus of languages resembles a similar quest that went on some 20 years ago when the assumption of an involvement of the non-dominant hemisphere in second language (L2) processing led to series of experiments aimed at elucidating this lateralization pattern in bilinguals. In a strongly voiced article, Paradis (1990) has tried to put an end to the ‘search for the Loch Ness Monster’, but this has been only partially successful. Something similar is now happening with the localization of languages. Since languages can be used independently, the assumption was and is that different languages have their own neural substrate. Attempts to show this seemed to be successful in the beginning (Kim *et al.*, 1997), but more recently the majority of studies on this issue conclude that there is hardly any evidence for separate substrates for different languages.

Like the case of the non-dominant hemisphere involvement, the focus has now shifted to finding the conditions under which differences in localization can be evidenced. In this vein Stowe (2006: 309) talks about:

two factors that increase the likelihood of finding a difference between L1 and L2: First a relatively difficult aspect of L2 processing should be tested (e.g. comprehension of complex sentences). Second, analysis techniques that focus on individuals rather than groups might also optimize the design.

So, L2-specific localizations are the exception rather than the rule now and can only be shown in very specific conditions. Interestingly, NI research does seem to have ended the debate on the involvement of the non-dominant hemisphere. On the basis of the recent literature on this and his own research, Indefrey (2006) concludes that the non-dominant hemisphere is not involved in processing other languages more than it is in processing the first language (L1).

Differences in localization that have been identified seem to be caused by task differences. Hasegawa *et al.* (2002: 657) have suggested

that such differences may reflect general cognitive load effects rather than L1 or L2 specific linguistic aspects:

The greater cortical recruitment for L2 may be at least partially a reflection of its greater demands, rather than reflecting cortical areas that are specifically dedicated to L2, although this possibility is not precluded.

A consequence of this line of thinking is that low fluency is associated with more effort and thus enhanced activation, and higher proficiency with less effort and thus reduced activation.

The problem with this solution is that there is no non-circular definition of cognitive load: there is no way we can assess cognitive load independent of techniques used. Slower reaction times as compared to control conditions are interpreted as more cognitive load. Similarly, more activation is not the same as cognitive load, but a correlate. In their discussion, Hasegawa *et al.* (2002: 657) add another interesting point that is not mentioned often: ‘Less activation does not always mean that the task is easy for the individual, but could mean that s/he is not attempting to meet the required demand.’ Now, it could be argued that the findings on localization have changed our view on multilingual processing, but the assumption of non-selective access has been around for quite some time now, based on more traditional experimental techniques such as interlingual priming and the neighbour effects in lexical access (Dijkstra, 2005). No new insights have been gained from the NI data on this issue.

IV Are the learners’ relevant background characteristics assessed adequately?

Age of acquisition, level of proficiency and language contact appear to be the main variables that are included in the research reported on. These factors are operationalized in different and mostly inadequate ways. It should be stressed here that this is not a problem for NI studies only. The majority of psycholinguistic studies on multilingual processing show the same weaknesses. For the assessment of language proficiency, a frequently used method is to ask participants to do a self-evaluation of their language skills. For example, Hasegawa *et al.* (2002) used self rating of listening fluency in L2 using a 5-point scale. One problem is that correlations between self-assessment and formal tests for language proficiency are notoriously low (for a large scale investigation including such measures,

see Bonnet, 2004). Another problem is that proficiency is presented as a mono-dimensional concept that can be measured with simple tests or self evaluation questions, while there is a range of skills that can and should be distinguished. In NI studies that do measure proficiency using formal tests, a wide range of techniques is used. Perani *et al.* (1998) used a translation task, while Klein *et al.* (1999) used a linguist's evaluation of spoken language and the vocabulary part of the WAIS-R test, a widely used test of components of intelligence. In some studies like Kim *et al.* (1997), no information on proficiency is given at all.

Obviously, the use of such a diversity of procedures makes a comparison between studies very problematic. In the European context, the development of the Common European Framework of Reference (see http://www.coe.int/t/dg4/linguistic/CADRE_EN.asp) has allowed researchers to develop language proficiency scales that make it possible to compare populations with different levels of proficiency and socio-cultural settings. An example of a study on fMRI in multilinguals in which language proficiency is tested on the basis of this Framework is Wattendorf *et al.* (submitted for publication). Tests based on this framework would make it easier to compare levels of proficiency between experiments. With the instruments used now, it is extremely difficult to make such comparisons.

Similarly, contact with the language is not assessed in even remotely appropriate ways in much of the NI research. Detailed information about contact is crucial because it will easily override the effects of age or way of acquisition. In sociolinguistics, there is a range of methods and instruments that have been used to estimate the different types and quantities of contact with the relevant languages. Type of language use (written/spoken), intensity and relevance of contact, complexity of the language used, language use in social networks: all these issues are potentially relevant. Exceptions are the studies reported on by Perani *et al.* (2003), who administered a detailed language use questionnaire covering several types of contact with both languages investigated, and Wattendorf *et al.* (submitted for publication), who also used an extensive sociolinguistic questionnaire. It would be very useful if a similar kind of assessment of contact would be used in different studies. Recently, a new questionnaire on language history has been developed at the University of Richmond, which may be useful for other researchers since there is both a long and a short

version that are manageable and sufficient for most types of psycholinguistic research (see http://cogsci.richmond.edu/questionnaire/L2_questionnaire.html).

Age of acquisition is also treated as a one-dimensional variable without taking into account relevant aspects of the acquisitional setting, such as the language of caretakers and peers, type of education in which different languages are used (bilingual education, immersion) and the amount of input of the various languages in these settings. Also, the intensity of contact in the beginning period may be a crucial variable. Starting at age five with one or two hours of foreign language education in primary education – as is now becoming customary in many European countries – is fundamentally different from a full early immersion approach as found in the Canadian context, while formally the age of acquisition in both settings may be similar.

At the same time, many variables that have been shown to play a role in second language development for some reason play no role in the selection of informants and set-up of experiments of both NI experiments and many behavioural studies. These factors include: motivation to learn the language, language aptitude, attitudes towards L1 and L2, level of L1 development, other languages learnt and degree of literacy, to name just the most obvious ones. To make matters worse, all these factors are likely to show a complex pattern of interaction over time, leading to highly individual patterns of development and resulting in considerable variation between and within individuals.

V Variation as a problem or a source of information?

There can be little doubt that the developmental path of L2 development is generally more variable among individuals than L1 development. This reflects the larger range of factors playing a role in L2 development. Variability in the input is probably the single most important factor in this difference between L1 and L2 development. Variation plays a role in different ways. Stowe (2006) and Uylings (2006) point to the problems related to inter-individual variation in brain anatomy. Group means become less relevant when there are large inter-individual differences in this respect. Indefrey (2006) points to the problem of the use of group means and the lack of interest in individual variation in

activation patterns. He refers to an electrical stimulation study by Lucas *et al.* (2004: 290), which shows that:

although there might not be any regions that are exclusively recruited or exclusively necessary for L1 or L2 word production *across* individuals there seem to be cortical sites in many individuals that are only necessary for word production in one of the languages.

One solution for the inter-individual variation problem is the use of intra-individual comparisons of the functional architecture for different languages. Examples of this approach can be found in Hasegawa *et al.* (2002) and Mahendra *et al.* (2003). In the latter study, early and late learners of L2 were compared using a number of tasks in L1 and L2. The outcomes are to a certain extent alarming: the degree of separation for regions of activation for L and L2 'did not exceed that associated with run-to-run variability for either task or either group' (Mahendra *et al.*, 2003: 1225).

Osterhout *et al.* (2006) also discuss some of the problems related to variability. They point out that in order to get the necessary signal-to-noise ratio, data have to be averaged over fairly large numbers of items first and then over subjects. L2 speakers will typically show more variation than monolinguals. This means that comparisons between monolinguals and bilinguals are hampered by the large variation in the latter group: similar effect sizes in the two groups may be significant in the more homogeneous monolingual group but may not reach significance because of the high variation in the bilingual group. Similarly, statistical analyses in which interactions with the group factor are included will be problematic. Osterhout *et al.* discuss an ERP (Event Related Potentials) study by Hahne (2001) on the perception of syntactic anomalies in monolingual Germans and Russian learners of German. The measure focused on was a left anterior negativity (LAN) effect which has been associated with morpho-syntactic violations. Though many of the individual learners showed a LAN effect as did the monolinguals, no significant effect was found for the learners due to variation. It is remarkable that Osterhout *et al.* (2006) argue for a focus on individual patterns, while in their own study they only present data on the group level.

One aspect that seems to have escaped attention so far is the impact on processing that use of a large number of items to solve the signal-to-noise ratio problem of NI techniques in itself may itself have. The processing of large numbers of items may have an impact on the processes studied.

In addition, averaging takes away the variation over time that may be informative about short term aspects of variation. In addition to means analysis, it may be useful to consider variation analysis as an additional way of analysing the data. Holden (2002: 57) argues that ‘a variability analysis is concerned with identifying systematic changes in variability; it assumes that the pattern of changes in variability is informative, that it may reflect the intrinsic dynamics of the system’. In fact, ERP data may lend themselves perfectly to developing a dynamic model of change over time because they are such dense data. A dynamic systems approach looks at the interaction of variables over time, acknowledging their fundamental dynamic rather than static nature. This approach has been shown to be very promising for the study of second language development (de Bot *et al.*, 2007) and holds a similar promise for application of NI data in the study of that process.

The research reported on in Gullberg and Indefrey seems to indicate that bilinguals show more variation than monolinguals. In the studies mentioned, variation is treated as a nuisance variable obscuring the ‘real’ differences. This approach is at variance with more recent approaches to variation in first and second language development (van Dijk and van Geert, 2005; Verspoor *et al.*, submitted for publication), which show that variation is both the outcome of and a necessary condition for development, and that it can inform us about the dynamics of the developmental process. This may hold for NI data as much as it does for behavioural and experimental data.

VI What does NI tell us about multilingual development?

One of the clearly missing issues in the research on second language development (SLD) and NI is the study of development over time. There is quite some research on the impact of different developmental settings on neural organization. In particular the early/late distinction has attracted considerable interest. As Birdsong’s (2006) contribution to the volume under review shows, the early/late distinction as a reflection of the existence of a critical period in SLD is a very complex one. The arguments for choosing a particular age range as ‘early’ are often lacking, apart from a vague reference to critical periods. An example is the Mahendra *et al.* (2003) study in which the early bilinguals have a mean age of one year at

the onset of acquisition with a range of zero to five, and the late bilinguals a mean age of acquisition of 11 with a range of 6 to 15. An early start does not mean that much in itself. When a language is not used after that early start, it may disappear completely as the work by Ventureyra (2005) has shown. She studied first language retention in France, in Koreans who had been adopted and moved to France between the ages of four and eight. Her data show that there are no signs of any retention of Korean, neither in behavioural tests data nor in brain activity. In other words, studying the early/late distinction can only be done when the rest of the language development history is taken into account.

Another issue is that studies carried out so far may tell us something about language proficiency at different ages or stages, but very little about the process of development itself. Even studies like the ones reported on by Indefrey (2006) and Osterhout *et al.* (2006) that use a longitudinal design provide information about moments in time, not about how the system moves from one state to the next. For this, the dynamic systems based approach that provides tools to study growth in complex adaptive systems may be fruitful (Schumann, 2006; de Bot *et al.*, 2007).

VII The ‘native imager’ issue

Another holy grail has been the search for native-like activation patterns in L2 learners (e.g. Perani *et al.*, 2003). Green *et al.* (2006: 100) argue that over time through language use the representations of L2 in L2 speakers will become more and more similar to those of native speakers of that language: ‘a natural expectation is that the neural representation of a second language (L2) will converge with that of the native speakers of that language as proficiency improves.’ Why this should be the case is not clear. Given the difference in initial conditions between monolinguals and bilinguals (apart from those who have multilingualism as their mother tongue) the emergence of native-like patterns of activation in bilinguals is actually extremely unlikely. The idea may be that for native-like proficiency, native-like activation patterns are required. But this is clearly not necessary. Multilinguals do not use their languages in the same way as monolinguals. They are multilingual because they use more than one language, and for the functions for which monolinguals have only one code, multilinguals have two or

more. Provided that they use multiple languages, the amount and type of use of each individual language will be different from that of monolingual native speakers. As Green *et al.* (2006) indicate, there may be degeneration, i.e. different processes and mechanisms can lead to the same (behavioural) results. The basic idea is that input and language use have considerable impact on brain activity. As mentioned earlier, Hagoort (2006: 94) also argues that variability in input will have an impact on functional specialization in the brain. Given the different developmental paths monolinguals and multilinguals follow in their development, as Grosjean (1998) has convincingly shown, it is very unlikely that patterns of activation will converge.

As Sabourin (2003) has argued, we need to look not only at whether L2 speakers can acquire the same knowledge as native speakers, but in particular how they use that knowledge. ERP studies may help us to better understand the use part, but whether they will also help us to understand the knowledge part is less clear.

VIII What is the added value of NI data?

One of the main issues in the current debate on NI and multilingualism is what the added value is, or, to put it differently: how should we interpret discrepancies between behavioural data, data from bilingual aphasia and NI data? Understandably, most authors will interpret a difference in findings between behavioural data and NI data as added value: the NI data reveal something that does not come out in behavioural data. NI techniques are more sensitive in detecting changes in cognitive processes. This line of thinking can be found in the contributions by Green *et al.* (2006) and Rodriguez-Fornells *et al.* (2006) to the volume under discussion.

A less optimistic perspective is that there is a discrepancy that reveals a gap between the two types of data, and that there is no real criterion to give a higher priority to NI data than to behavioural data. The question is always what the data are supposed to reflect. If we are interested in language behaviour, then behavioural data are to be preferred. But we are typically not interested in the behaviour, but in the processing mechanisms behind it. Somehow we feel that NI data are closer to what we really want to know about underlying processes than behavioural data. To what extent is this assumption warranted? Is the difference between the two types of

data to be explained by the processing that takes place somewhere in between the neuro-physiological processes that we measure with NI techniques and the behaviour that we measure with techniques such as lexical decision and sentence interpretation? Should these in-between processes be seen as noise-creating steps, or are they essentially part of the same processes we are interested in? Despite the progress that has been made in linking the functional and the neuro-level (see Indefrey, 2006; 2007), we do not know what it is that we measure with NI techniques. There is a signal, but what is the source? Is the level of processing we measure with NI techniques (as far as we can say that they do measure on the same level) just another level with its own peculiarities, or is it the next layer that brings us closer to the real source or locus of language processing? I assume that most NI researchers would at least hope that the latter is true, but the problem that arises when taking this perspective is that we encapsulate the language system and reduce it further and further to a collection of processes deep in the brain. This perspective on cognition is totally at odds with the now-prevalent view of cognition as shared, situated and embodied (Beer, 2000).

IX The imaging of what?

The discussion of how different types of data are related brings us to the main question to be answered: what is it that is represented in NI? There is considerable discussion on what NI can tell us about language processing. Most contributors to the Gullberg and Indefrey volume seem to have few qualms about this. For instance, Mueller (2006) is fairly clear about the interpretation of different NI phenomena: 'Learners' ERP responses reflect a highly strategic use of case marking postpositions' (p. 263) and 'the finding of ERP components that indicate relatively automatic syntactic processes, namely anterior negativities' (p. 247). She also argues that absence of a particular NI effect is proof of certain processes: 'Nonnative participants seemed to underuse automatic syntactic processes, indicated by the deviant topographical distribution of the early negativity, as well as thematic processes, as reflected in the absence of the effect' (p. 255). This is problematic. Even if a correlation between the instantiation of an effect at the neural level and at the functional level can be shown, this does not mean that the absence of a specific NI effect implies that the associated

functional process does not take place. For example, the absence of a P600 effect, one of the better known ERP components, is no proof that a sentence is interpreted as being non-deviant. Paradis (2004: 154) points out that techniques may not be sensitive enough to detect activation: ‘The fact that something is *not* detected is no indication that a particular area is not active: only that the technique employed does not pick it up (either because it is below the sensors’ threshold or because it is masked by the baseline task)’. The point mentioned by Hasegawa *et al.* (2002) on lack of attention to task discussed earlier may be relevant here as well.

As Mueller (2006) indicates, ERP effects have been found for non-linguistic information processing, which might suggest general cognitive processes and not specifically linguistic ones. The argument that some effects, like LAN, have not been reported for non-linguistic information processing cannot be used to say that therefore other ERP effects are linguistic in nature. There are, no doubt, specific ERP effects in other types of information processing such as visual perception, that have no equivalent in language processing. An example is the visual mismatch negativity (VMMN) component reported on by Czigler *et al.* (2006: 4). The editors of the book are considerably more cautious on this issue:

It is also problematic that, at least in the language domain, we lack a real understanding of the functional significance of differences at the neural level. Observed differences tend to be interpreted in a circular manner starting from the (plausible) assumption that whatever is found in the more proficient speakers must be more effective, be it an increase or decrease of hemodynamic activation, an increase or decrease of gray matter.

The correlation between language tasks and ERPs has been interpreted as a sign that some type of linguistic activity takes place in the areas showing these potentials, but it may well be that some non-linguistic effect, such as unexpectedness, causes the brain activity.

Schumann (2006; personal communication) argues for more attention for the underlying connectivity of different language related areas. He maintains that:

if NI indicates activity in a certain part of the brain in response to some second-language behaviour, without knowledge of the afferent and efferent connections to that area and without knowledge of what other functions that area subserves, interpretations of the activity will be very difficult to make. (personal communication)

In other words, research on multilingual processing should be based on fairly deep neuro-biological knowledge rather than a sketchy set of notions of areas and their assumed functions.

X The contribution of NI research to theories about the multilingual brain

In their evaluations of the contribution of NI to theories in cognitive psychology, Coltheart (2006) and Page (2006) ask whether NI data can help to decide between competing theories. Examples they give are serial vs. parallel processing in reading and the semantic contributions to reading aloud. For both issues there are clear predictions on the behavioural level and there is substantial evidence to support one or the other position. Coltheart and Page argue that none of the NI studies that are devoted to these topics have provided convincing support for any of the positions on these theoretical issues. Uttal (2001: 217) rather bluntly indicated why that is the case:

Even if we could associate precisely defined cognitive functions in particular areas of the brain (and this seems highly unlikely), it would tell us very little if anything about how the brain computes, represents, encodes, or instantiates psychological processes.

The approach Coltheart and Page follow is this: they select a theoretical issue on which there are two markedly different positions that have been defended on the basis of functional behavioural data, then they formulate them as precisely as possible as testable hypotheses and indicate what NI data could refute or support either position in the debate.

It may be useful to follow this procedure and see whether a similar approach might work for multilingual processing. One of the core issues in multilingual processing in recent years has been the selective/non-selective access debate. Selective access means that in language production and perception in multilinguals, only the lexical elements from one language are activated, while non-selective access means that elements from more than one language are activated (Laheij, 2005). Following Coltheart and Page:

T_a (selective access): Only lexical elements from one language are activated.

T_b (non-selective access): Lexical elements from more than one language are activated.

What predictions follow with respect to NI data? Are there data that allow us to choose between T_a and T_b ? First, it would have to be assumed that languages have their own neural substrates. The second step would be to carry out experiments on lexical access with picture naming to see, whether along with the target word in the target language, words from other languages will be activated.

Several experiments using picture/word interference paradigms have shown that the non-selective position is supported (Hermans *et al.*, 1998; Laheij, 2005). There is hardly any NI research that aims at solving this issue: Rodriguez-Fornells *et al.* (2006: 137) list numerous studies showing that languages have overlapping neuro-anatomical representations, but seem to cling to a smaller set of studies ‘that are clearly in favor of a partial functional separation of bilingual lexicons in the brain’. The studies referred to are Dehaene *et al.* (1997) and Kim *et al.* (1997). Dehaene *et al.* (1997: 3809) had their participants listen to spoken text in L1 and L2 and their data showed that:

[i]n all subjects, listening to L1 always activated a similar set of areas in the left temporal lobe, clustered along the left superior temporal sulcus. Listening to L2, however, activated a highly variable network of left and right temporal and frontal areas, sometimes restricted only to right-hemispheric regions.

How this relates to ‘partial functional separation of the bilingual lexicon’ is not clear. Kim *et al.* (1997: 171) tested early and late bilinguals with a silent speech task with fMRI. They conclude that:

within the frontal-lobe language-sensitive regions (Broca’s area), second languages acquired in adulthood (‘late’ bilingual subjects) are spatially separated from native languages. However, when acquired during the early language acquisition stage of development (‘early’ bilingual subjects), native and second languages tend to be represented in common frontal cortical areas.

Apart from the fact that subsequent experiments failed to reproduce this effect, it is again not clear what these findings tell us about the lexicon. It seems that the conclusion drawn by Rodriguez-Fornells *et al.* is very indirect at best and clearly overoptimistic with respect to the evidence NI data can provide about a core issue like the selective/non-selective one.

XI Concluding remarks

In his book *A neurolinguistic theory of bilingualism*, Michel Paradis (2004) devotes a whole chapter to NI studies of the bilingual brain. He raises a large number of pertinent issues that cast doubt on the validity of NI studies for our understanding of language processing in general, and the study of the multilingual brain in particular. The conclusions we can draw on the basis of the articles presented in the Gullberg and Indefrey book seem to be in line with Paradis’s main points. The validity of the procedures used, contradictory findings using different techniques, partial divergence of NI studies and studies based on lesions and, most

importantly, what it is that we measure with NI techniques all support Paradis's conclusions. There may be progress within the subfield as defined by the members of the community, but for the relative outsider a number of fundamental problems need to be solved before the contributions on NI can be added as valid new knowledge that enhances our understanding of the multilingual brain. The success of the NI enterprise seems to have gone at the expense of some of the basic requirements of empirical research: validity, convergence, reliability and replicability. The NI community has gone off on its own, but it might still learn from the insights and skills of researchers who have been working in the field of multilingualism for longer than the last two years or so.

This rather negative evaluation of the contribution of NI to understanding cognitive aspects of multilingualism reflects a more general concern expressed by several cognitive scientists who conclude that NI so far has not contributed to theory by providing data that can be used to distinguish between competing psychological theories (Coltheart, 2006; Page, 2006). Some even argue that studying the brain will tell us nothing about cognition:

The essence of these arguments is that distinct neurological structures need not correspond to functional modules – indeed, there might not be any modules. To be able to decide whether there are and whether there is any correspondence, you need to have a complete theory of cognition before you begin interpreting images. Hence imaging can, in principle, add nothing new. There is a level of psychological theorizing – the cognitive level – that can only be studied at this level, and information from lower levels will tell us nothing about what happens at the cognitive level. (Harley, 2004: 10–11)

In his discussion of the contribution of functional neuro-imaging to cognitive psychology, Page (2006) mentions a number of additional arguments why NI can be seen as a threat to cognitive psychology as a field of work and research. First, NI is very expensive. Data from one single participant may well cost 100 times more than in conventional experiments. Second, 'large capital investments tend to concentrate subsequent funding around them . . . Indeed, [funding agencies] may even be tempted to provide lavish funds in an attempt to justify their original decision' (Page, 2006: 14). Money supply for research is limited and the costs of NI detract from other types of research while the results of NI in terms of scientific progress do not justify such a disproportionate spending. Third, NI research leads to a brain drain of talented students and post docs who understandably go where the money is.

Page's final two arguments have to do with public relations and the image of the field for the larger public. NI provides the kind of pictures the media are fond of. The presentation of research results suggests that NI is all cognitive psychology does and that there can therefore be no really good research that does not look inside the brain. Finally, he argues that the application of NI may lead to a medicalization of cognitive psychology. Research involving NI tends to be dressed up as medical research. Is there a risk that the medical community will take over the research agenda and that research only gets funded when some medical application is promised?

Not all of Page's arguments apply equally to the field of the psycholinguistics of multilingualism, but a considerable part of our research has strong links with cognitive psychology, as evidenced by the many reports on multilingual processing in leading journals in that field. As I have argued, the return on investment for NI research on multilingual processing so far is low. In that sense it supports Page's position. Both researchers and funders may want to consider to what extent an increase of NI research is warranted. In many countries large facilities for NI have been set up in the last decade and many of them will be reviewed in the next few years. These assessments may well lead to a shake-out based on the type of arguments presented here.

Acknowledgements

The author is indebted to John Hoeks, Peter Indefrey, Ludmila Isurin, Wander Lowie, John Schumann, Margaret Thomas, Marjolijn Verspoor and an anonymous *Second Language Research* reviewer for their comments on an earlier version of this contribution, and to Sybrine Bultena for her support in finalizing the manuscript.

XII References

- Beer, R.** 2000: Dynamical approaches to cognitive science. *Trends in Cognitive Sciences* 4, 91–99.
- Birdsong, D.** 2006: Age and second language acquisition and processing: a selective overview. In Gullberg, M. and Indefrey, P., editors, *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell, 9–49.

- Bonnet, G.** 2004: *The assessment of pupils' skills in English in eight European countries 2002*. Paris: European Network of policy makers for the evaluation of education systems.
- Coltheart, M.** 2006: What has functional neuroimaging told us about the mind (so far)? *Cortex* 42, 323–331.
- Czigler, I., Winkler, I., Pato, L., Varnagy, A., Weisz, J. and Balazs, L.** 2006: Visual temporal window of integration as revealed by the visual mismatch negativity event-related potential to stimulus omissions. *Brain Research* 1104, 129–40.
- de Bot, K., Verspoor, M. and Lowie, W.** 2007: A dynamic systems theory approach to second language acquisition. *Bilingualism, Language and Cognition* 10, 7–21.
- Dehaene, D., Dupoux, E., Mehler, J., Cohen, L., Paulescu, E. and Perani, D.** 1997: Anatomical variability in the cortical representation of first and second languages. *NeuroReport* 8, 3809–15.
- Dijkstra, T.** 2005: Bilingual visual word recognition and access. In Kroll, J. and de Groot, A., editors, *Handbook of bilingualism: psycholinguistic approaches*. Oxford: Oxford University Press, 179–201.
- Dijkstra, T. and van Heuven, W.** 2006: On language and the Brain: or on (Psycho)linguistics and Neuroscientists? Commentary on Rodriguez-Fornells *et al.* In Gullberg, M. and Indefrey, P., editors, *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell, 191–98.
- Green, D., Crinion, J. and Price, C.** 2006: Convergence, degeneracy and control. In Gullberg, M. and Indefrey, P., editors, *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell, 99–125.
- Grosjean, F.** 1998: Studying bilinguals: methodological and conceptual issues. *Bilingualism: Language and Cognition* 1, 131–49.
- Gullberg, M. and Indefrey, P.** 2006: *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell.
- Hagoort, P.** 2006: What we cannot learn from neuro-anatomy about language learning and language processing: a commentary on Uylings. In Gullberg, M. and Indefrey, P., editors, *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell, 91–97.
- Hahne, A.** 2001: What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research* 30, 251–66.
- Harley, T.** 2004: Promises, promises. *Cognitive Neuropsychology* 21, 51–56.
- Hasegawa, M., Carpenter, P. and Just, M.** 2002: An fMRI study of bilingual sentence comprehension and workload. *Neuroimage* 15, 647–60.
- Hermans, D., Bongaerts, T., de Bot, K. and Schreuder, R.** 1998: Producing words in a foreign language: can speakers prevent interference from their first language? *Bilingualism: Language and Cognition* 1, 213–29.

- Holden, J.** 2002: Fractal characteristics of response time variability. *Ecological Psychology* 14, 53–86.
- Indefrey, P.** 2006: A meta-analysis of hemodynamic studies on first and second language processing: which suggested differences can we trust and what do they mean? In Gullberg, M. and Indefrey, P., editors, *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell, 279–304.
- 2007: Brain imaging studies of language production. In Gaskell, G., editor, *Oxford handbook of psycholinguistics*. Oxford: Oxford University Press, 547–64.
- Kim, K.H.S., Relkin, N.R., Lee, K.M. and Hirsch, J.** 1997: Distinct cortical areas associated with native and second languages. *Nature* 388, 171–74.
- Klein, D., Milner, B., Zatorre, R., Zhao, V. and Nikelski, J.** 1999: Cerebral organization in bilinguals: a PET study of Chinese–English verb generation. *NeuroReport* 10, 2841–46.
- Klein, W.** 1989: Introspection into what? Review of Faerch, C. and Kasper, G., editors, *Introspection in second language research 1987*. *Contemporary Psychology: A Journal of Reviews* 34, 1119–20.
- Koopmans, P.** 2006: L2 acquisition, age, and generativist reasoning: a commentary on Birdsong. In Gullberg, M. and Indefrey, P., editors, *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell, 51–58.
- Laheij, W.** 2005: Selection processes in monolingual and bilingual lexical access. In Kroll, J. and de Groot, A., editors, *Handbook of bilingualism*. Oxford: Oxford University Press, 289–307.
- Lucas, T., McKahn, G. and Ojemann, G.** 2004: Functional separation of languages in the bilingual brain: a comparison of electrical stimulation language mapping in 25 bilingual patients and 117 monolingual control patients. *Journal of Neurosurgery* 101, 449–57.
- Mahendra, N., Plante, E., Magloire, J., Milman, L. and Trouard, T.** 2003: fMRI variability and the localization of languages in the bilingual brain. *NeuroReport* 14, 1225–28.
- Mueller, J.** 2006: L2 in a nutshell: the investigation of second language processing in the miniature language model. In Gullberg, M. and Indefrey, P., editors, *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell, 235–70.
- Osterhout, L., MacLaughlin, J., Pitkänen, I., Frenck-Mestre, C. and Molinaro, N.** 2006: Novice learners, longitudinal designs and event-related potentials: a means for exploring the neurocognition of second language processing. In Gullberg, M. and Indefrey, P., editors, *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell, 199–230.
- Page, M.** 2006: What can't functional neuroimaging tell the cognitive psychologist. *Cortex* 42, 428–43. Also published online.
- Paradis, M.** 1990: Language lateralization in bilinguals: enough already! *Brain and Language* 39, 576–86.

- 2004: *A neurolinguistic theory of bilingualism*. Amsterdam: John Benjamins.
- Perani, D., Abutalebi, J., Paulescu, E., Brambati, S., Scifo, P. and Cappa, S.** 2003: The role of age of acquisition and language usage in early, high-proficient bilinguals: an fMRI study during verbal fluency. *Human Brain Mapping* 19, 170–82.
- Perani, D., Paulesu, E., Sebastian-Galles, N., Dupoux, E., Dehaene, S., Bettinardi, V., Cappa, S.F., Fazio, F. and Mehler, J.** 1998: The bilingual brain: proficiency and age of acquisition of the second language. *Brain* 121, 1841–52.
- Rodriguez-Fornells, A., de Diego Balaguer, R. and Münte, T.** 2006: Executive control in bilingual language processing. In Gullberg, M. and Indefrey, P., editors, *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell, 133–90.
- Sabourin, L.** 2003: Grammatical gender and second language processing: an ERP study. Unpublished doctoral dissertation, University of Groningen, Groningen.
- Schumann, J.** 2006: Summing up: some themes in the cognitive neuroscience of second language acquisition. In Gullberg, M. and Indefrey, P., editors, *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell, 313–19.
- Stowe, L.** 2006: When does the neurological basis of first and second language processing differ? Commentary on Indefrey. In Gullberg, M. and Indefrey, P., editors, *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell, 305–11.
- Stowe, L., Haverkort, M. and Zwarts, F.** 2005: Rethinking the neurological basis of language. *Lingua* 115, 997–1042.
- Uttal, W.** 2001: *The new phrenology: the limits of localizing cognitive processes in the brain*. Cambridge, MA: MIT Press.
- Uylings, H.** 2006: Development of the human cortex and the concept of ‘critical’ or ‘sensitive’ periods. In Gullberg, M. and Indefrey, P., editors, *The cognitive neuroscience of second language acquisition*. Malden, MA: Blackwell, 59–90.
- van Dijk, M. and van Geert, P.** 2005: Disentangling behavior in early child development: interpretability of early child language and the problem of filler syllables and growing utterance length. *Infant Behavior and Development* 28, 99–117.
- Van Orden, G. and Paap, K.** 1997: Functional neuroimages fail to discover pieces of mind in the parts of the brain. *Philosophy of Science* 64, S85–S94.
- Ventureyra, V.** 2005: A la recherche de la langue perdue: étude psycholinguistique de l’attrition de la première langue chez des Coréens adoptés en France [In search of the lost language: a psycholinguistic study of first language attrition among Korean adoptees in France]. Unpublished

PhD Thesis, Ecole de Hautes Etudes en Sciences Sociales (EHESS), Paris.

Verspoor, M., Lowie, W. and van Dijk, M. forthcoming: Variation in second language development: a DST Perspective. *The Modern Language Journal*.

Wattendorf, E., Westermann, B., Keil, U., Zappatore, D., Franceschini, R., Lüdi, G., Radue, G., Rager, E. and Nitsch, C. submitted for publication: Early second language acquisition modulates the cortical pattern of language-related activity: an fMRI study of multilinguals.