The bilingual brain

Michaël Mouthon, Jean-Marie Annoni, Asaid Khateb

Laboratory for Cognitive and Neurological Sciences, University of Fribourg, Faculty of Sciences, Section of Medicine, Switzerland

Neurology Unit, Hospital of Fribourg, Switzerland

The Laboratory for the Study of Bilingualism, Edmond J. Safra Brain Research Center for the Study of Learning Disabilities and Dept of Learning Disabilities, Faculty of Education, University of Haifa, Israel

Acknowledgement: This work was partially supported by the Swiss National Science Foundation (grant no’ 325130-138497 to J.M. Annoni) and by the Israeli National Science Foundation (Grant no 623/11 to A. Khateb). Part of this text is based on the PhD dissertation of Dr. M. Mouthon.

Summary

Bilingualism can be defined simply as the regular use of two languages and is a phenomenon that affects almost every nation in the world, but finds a natural geographical context in Switzerland. At the descriptive level, bilingualism can vary in terms of how two languages are learnt and how proficient a bilingual is in the second (L2) language compared with the first one. We review data which suggest that the different language areas, although a greater convergence is found when L2 proficiency increases, that the majority of bilinguals’ aphasic symptoms can be explained by deficits in languages control mechanisms and that (3.) language control is subserved by a left lateralised fronto-subcortical circuit which is at least partially within the general cognitive control system.

Key words: bilingualism; language representation; language selection

Bilingualism: definitions and major research issue

The existence of different geographical and ethnic groups in the world led to the development of a wide variety of languages, a word (i.e. “language”) that defines the set of modalities that humans use to communicate with each other. The development of communication between these groups and globalisation processes (immigration, culture, politics, tourism …) highlighted the special ability of humans to learn and speak several languages. This special ability is commonly referred to as “bilingualism” and has a crucial role in the construction of our societies, since bilingualism is a socio-linguistic phenomenon that affects almost every nation in the world, all classes of a society and all its age groups. Currently, more than half of the population of the world speaks more than one language and, because of the disappearance of the borders between certain countries in some parts of the world (as for instance in Europe), the need to learn more and more “foreign” languages increases.

At the conceptual level, several definitions exist for what the phenomenon bilingualism represents. A restrictive definition already offered in 1935 by Bloomfield [1] proposes that bilingualism implies a simultaneous learning of two languages during childhood, where a similar proficiency is reached in both. A less conservative view defines a bilingual as a person who manages at least one of the three major linguistic abilities (i.e., either comprehension or speaking or writing) in a language other than the mother tongue [2]. Finally, one of the most commonly used definitions sees bilingualism as the ability to express oneself in a mother tongue (called L1) and at least in a second learned language (called L2) in daily life [3]. This ability allows bilinguals to understand and express themselves through an oral or written modality in both languages acquired, independent of the level of proficiency in L2. Other extended definitions of bilingualism take into consideration the age of acquisition of L2, such that: (1.) “early bilingualism” concerns people who learnt their L2 before 5–9 years old, (2.) “late bilingualism” refers to people who learnt after this period and (3.) “simultaneous bilingualism” or “natural bilingualism” defines those people who went from speaking no language to speaking using two languages simultaneously from the first day [4, 5].

In the context of bilingualism, it is important to distinguish between a bilingual person and a bilingual behaviour, which is determined by the communicative context in which a bilingual person expresses himself. Thus, Grosjean [3] distinguished two types of bilinguals’ behaviour, depending on whether they speak to a monolingual or to another bilingual person. In the monolingual mode, the bilinguals chose one language and deactivate (or inhibit totally) the other [6], thus forcing themselves to stay in the language of the interlocutor. In the bilingual contexts, they chose one language of reference as a function of a preferred language (generally of the listener) and keep high the activation level of the other as a function of the need, such that if the interlocutors have the same L1 and L2, they mix both languages and switch from one to another.

Correspondence:
Professor Asaid Khateb, Ph D
Edmond J. Safra Brain Research Center for the Study of Learning Disabilities and Dept of Learning Disabilities, Faculty of Education, University of Haifa
Mount Carmel
IL-Haifa 31905
Israel
Akhateb(at)edu.huifa.ac.il
Language representation in the brain

Since the seminal clinical works of Paul Broca and Carl Wernicke [7, 8], it is well accepted that language is localised in some specific areas of the brain that are lateralised in the left hemisphere. Since these studies, a fundamental distinction in language research was made between types of linguistic knowledge acquired through life, and the questions of how the orthographical, phonological, morphological, semantic, syntactic and grammatical information are organised and processed continues to occupy many researchers.

The “mental lexicon”, for instance, contains all memorised information about words, including their sound (phonetic), spelling (orthographic) and their meaning (semantics). A second type of knowledge is the “mental grammar”, which contains rules and operations that permit the combination of lexical forms and abstract representation into complex linguistic structures as sentences (i.e., syntax). Theoretical models, based on neurolinguistic and psycholinguistic research, propose that these two types of knowledge are managed by two memory systems [10–12]. The lexical memory depends on declarative memory processes, mainly rooted in temporal lobe structures (the hippocampus and related surrounding regions). This memory has been implicated in learning, representation, and use of semantic and events knowledge. The grammatical memory depends more on the procedural memory, which is rooted in left frontal-basal-ganglia circuits. This type of memory has been implicated in the unconscious learning and execution of long-established motor plans, skills or habits.

In the case of bilingual people, it has been suggested that the representations of L2 knowledge should differ in comparison with those of L1, particularly in late bilinguals [10]. Mainly, the grammatical knowledge and processing that is usually associated with the procedural memory in L1 would be more dependent on the declarative memory for L2. This shift from the procedural to the declarative memory is thought to be related, for example, to the conscious (explicit) learning, retrieval and use of nonautomatic grammatical rules (as in a pedagogic environment) or memorising in lexicon preexisting grammatical structures as expressions, sentences, idioms, etc. Thus, it was suggested, on the basis of these theoretical considerations, that while certain aspects of language representation in the bilingual brain are shared by L1 and L2, the dependence on declarative memory of certain other aspects of L2, which are affected by the age of acquisition and extent of exposure of the second language, would lead to different brain representations for these language aspects. In such a case, the shift should be more important for late learners than early learners, since the latter are thought to develop a certain automaticity with practice and age.

In the following section, we consider evidence for divergent and convergent language representations in bilinguals’ brains. The issue of the neural basis of bilingualism has attracted many researchers during the past two or three decades. In particular, the questions of how the brain organises and uses the different languages were the focus of multiple scientific clinical reports and functional neuroimaging investigations in bilinguals. We first review some evidence on how the brain organises the different languages acquired and on whether or not these share the same neural substrates? Some clinical and experimental data tended to indicate that L1 and L2 are represented differently in the left cerebral hemisphere. Some research even tended to suggest that, whereas language is left lateralised in monolinguals, in bilinguals language is represented bilaterally in the brain. In the later sections, we present some of the evidence that supported the divergence of representation for the different languages in the brain and then evidence for the convergence of representation in the light of recent functional brain imaging data.

Divergence of representations

Some clinical evidence: The question of language representation in the brain was originally raised on the basis of clinical observations of bilingual aphasic patients [13–22]. Several of these reports showed an unequal pattern of impairment in the two languages (selective impairment), whereas others suggested an unequal pattern of recovery of the two languages (selective recovery). Such observations led to the proposition that each language is represented in distinct cerebral areas and thus rely on different neural networks. For example, in one study of four bilingual aphasic patients with lesions limited to the left basal ganglia, it was reported that grammatical performance was more affected in the dominant than in the nondominant language [23]. It should, however, be noted that a previous anatomo-clinical analysis had failed to determine definitely the specific areas in the bilingual brain that participate in the processing of only one or the other language [24]. As for the hypothesis of bilateral language representation, clinical studies have shown that the incidence of aphasia with a right hemisphere lesion (in what is commonly called “crossed aphasia”, see [25]) in bilingual subjects is similar to that observed in monolinguals, thereby infirming the assumption of a different representation of languages in the first [26].
Some experimental evidence: In line with certain clinical observations that suggested that the different languages might rely on distinct neural substrates, electrophysiological brain stimulation studies also provided evidence for distinct representation of the different languages in the brain. Thus, various reports indicated that the stimulation of different cerebral locations, during neurosurgery monitoring, interfered unequally with the patients’ languages (27–34) (for a review see [35]). These data suggested that some cortical sites might selectively be involved in one of the languages but not the others. In particular, the effects of the stimulations were observed in the classically language-sensitive areas of the posterior temporal gyrus, and frontal and parietal areas, where stimulation induced deficits (decrease in production fluency, speech arrest, etc.) during language tasks such as object naming and reading. Thereafter, the development of functional brain imaging techniques (positron emission tomography [PET] and functional magnetic resonance imaging [fMRI]) provided new powerful tools for assessing this question of language representation. For example, some functional neuroimaging investigations reported different brain activation between L1 and L2, with different experimental paradigms in regions such as the Broca’s area, cerebellum, supramarginal gyri and others (e.g. [36–40]).

Convergence of representations

The alternative possibility that the different languages converge into the same language brain regions and that linguistic knowledge of L1 and L2 are processed by the same neural network finds support in both clinical and experimental data [41]. The theory of convergence proposes that the acquisition of L2 even in late learners involves the same brain network used for the L1 as in monolingual speakers, including the processing of semantic, syntactic, phonological but also grammatical information (for review see [42]).

Some clinical evidence: A common language representation for the different languages in bilinguals was supported by other clinical studies that reported similar impairments in both languages after brain damage (and a parallel recovery) [43–45]. For instance, in a very interesting case study conducted by Marangolo et al. [46], it was reported that both languages of a bilingual woman with chronic aphasia recovered after rehabilitation therapy in L2, a finding strongly suggesting that both languages relied on common neural substrates.

Some experimental evidence: Contrary to some functional studies that supported the hypothesis of a differential brain representation for L1 and L2, other studies suggested that the first and second use the same neural circuits[47–50]. For instance, support for the language representation convergence theory was reported in picture naming [51, 52] and semantic judgement tasks [48], where overlapping brain activation was observed for L1 and L2. Also, similar brain activations were observed during grammatical processing in early bilinguals [53]. However, the idea that both languages use a common network does not necessarily mean that the processing demands for the different languages will be identical. Actually, differences in the neural recruitment could be observed between languages in this network, probably owing to a “preference” or better performance for L1 than for L2 [54]. Hence, some evidence presently indicates that L1 and L2 rely on a shared network whose activation level might vary as a function of several modulating factors, including the level of proficiency in L2, age of acquisition and the extent of exposure to L2 [47–49, 55–58].

To summarise, two main visions about the issue of language representation in the brain prevailed in language research on bilinguals. According to the first, which was mainly supported by some clinical observations, electrophysiological stimulations and early functional imagery findings, the different languages were represented at least partially in different brain regions. The second vision proposed, thanks to the development of more rigorous techniques, analysis methods and controlled language paradigms, that the two languages share a common neural network. Nowadays, the discrepancy between the data provided by the different fields of research or from different studies from the same field (e.g., functional imaging research) are thought to result either from factors inherent to the studied populations (low vs high proficiency, early vs late, or balanced vs unbalanced bilinguals) or from a bad definition of the question addressed. Actually, a high number of functional results now support the hypothesis of a common brain network for both languages and document effects of proficiency and age of acquisition of L2 (see below). On the other hand, some authors propose that some of the clinical manifestations observed in bilingual aphasia, such as the presence of different recovery/impairment patterns that affected one language or the other language, do not attest to the presence of different language representation, but are linked to damage to the brain’s language control system (see below and see [41, 54, 59]).

Determining factors: age of acquisition, proficiency and exposure

The hypothesis that the age of acquisition of a new language might influence its representation had already been evoked by several early scientists who proposed the existence of a critical time period (between infancy and puberty) for language acquisition [60, 61]. In this regard, various data indicated that proficiency in grammar and pronunciation was less good with late acquisition age (e.g., [53, 62–64]). Such a critical period would reflect maturational changes and differences in the brain processes engaged. It has been suggested that learning processes were largely limited to implicit memory for children and to explicit memory for adults [10, 65], and that language learning abilities decrease over maturation owing to memory limitation [66, 67].

The effect of age has been also proposed to influence language lateralisation in the bilingual brain. Thus, some studies reported that the lateralisation of L2 could be influenced by factors such as the age of acquisition of L2, being more left-lateralised with early acquisition and more bilateral for late-acquired L2 [13, 68–70]. As already indicated, such a hypothesis was not verified in crossed aphasia studies and the effect of age of acquisition on language lateralisation and
representation remains a little unresolved since a recent meta-analysis (of 66 behavioural studies) concluded that there is bilateral lateralisation of both languages with early acquisition and left lateralisation with late acquisition [71].

From the functional neuroimaging perspective, various researchers have suggested that the age of acquisition of L2 is a critical factor for the functional organisation of the brain [40]. Kim et al. [40], for instance, reported that the two languages involved two different regions in Broca’s area in late bilinguals and a common region in early bilinguals. However, other studies suggested that language proficiency in L2 and exposure to this language have a much more determinant role in language representation than the age of acquisition [48, 55, 72] or the order of their acquisition [73, 74]. Thus, various results in the literature suggest that the brain’s representation of L1 and L2 converged to the same network when proficiency increased [75]. In one study [50], the comparison of functional responses in late and early bilinguals during a sentence comprehension paradigm showed that, independent of L2 age of acquisition, the two languages (L2 and L1) activated similar brain areas in highly proficient bilinguals, whereas in low proficiency subjects, L2-induced activation was more extensive. Other studies showed that brain activation for the two languages were similar when proficiency was comparable between L2 and L1 (e.g. [49, 51, 52, 76–78]), whereas low proficiency in L2 engaged additional brain activity (e.g. [47, 56, 79–81]). In such cases, it was suggested that differences between L1 and L2 were linked to the unequal computational demands that dynamically change with L2 proficiency [82]. As for exposure to the language, it has been found that L2 could replace L1 if bilinguals were no longer exposed to their L1, as shown with an adopted Korean population in France [83]. Another representative example of this switch in dominance was found in oral comprehension of Spanish-Catalan early bilinguals who had difficulties in distinguishing mispronounced and properly pronounced Catalan words (even if they had a native proficiency), because their exposition to Spanish was greater in their every-day life [84]. Taken together, many factors are thought to affect brain activation during L2 processing, and the possible differences observed between L2 and L1 appear to be the results of a dynamic interplay between the age of acquisition and exposure, which both act to determine the level of proficiency and the need or not for additional computational demands to perform the same tasks in the two languages.

Language control and selection mechanisms
in the brain

In a second line of research on the bilingual brain, the question of how the bilingual brain manages (or controls) the use of the two/or more languages was raised. Concretely, this question is related to the so-called “language control” or “language selection” mechanisms that allow bilinguals to communicate in one language rather than in the other, but also to switch from one language to another during the same conversation. This ability was the focus of multiple clinical and experimental studies for more than one decade. As already mentioned, some of the clinical manifestations of aphasia in bilinguals, such as the selective impairment and recovery of one language (the latter referred to as pathological fixation on one language) rather than the other in aphasic patients after brain injury, are thought by some authors to result from damage to the brain’s language control system [41, 54, 59]. Also, phenomena such as pathological language switching (defined as the uncontrolled alternation of utterances from one language to the other across the sentence boundaries) and language mixing (defined as the uncontrolled mixing of elements from the two languages within a single utterance), found after damage to the left prefrontal cortex [85] or inferior parietal cortex [86, 87] are also linked to this control system. Intuitively, if we assume that the different languages share a common neural basis, then it is also reasonable to assume that a control mechanism is needed to manage their use correctly and prevent interference from one to another during lexical selection/language production. Two major theories have been developed in the literature about this issue.

The late selection theory: This theory suggests that the lexicons of both the target and nontarget language stay active [88]. In this case, the competition occurs between lexical nodes or lemma to select the target language [89]. Selection mechanisms pick out the most highly activated lexical node at a given moment, and the degree of activation from the nontarget one will affect the ease of this selection (competition). In this respect, the lexical selection is achieved without need of active inhibition of the nontarget lexicon [90]. The greatest support for this view comes from the work of van Heuven et al. [91], which showed that the number of orthographic neighbours (e.g. PORK and WORK) in target and nontarget language interfered with the target word processing in L1 or L2. The emergence of words from a nontarget language during language production depends on L2 proficiency and is proposed to support this selection theory [92–94].

The early selection theory: In this theory, it is proposed that a prior inhibition of the nontarget language is necessary to solve the competition between lemmas of the different languages [95–98]. More precisely, the nontarget language is actively inhibited in the early stages of processing before access to lexicons [97] and only the target lexicon will stay active. Support for the early selection theory comes from clinical observations. For instance, Paradis et al. [99] reported the case of two bilingual aphasic patients who showed a paradoxical pattern between translation and spontaneous speech production. After an accident, they were able to speak in L1 but not to translate into it. At the same time, they were able to translate into L2 but not to speak it. The following day the same patients exhibited the opposite pattern. The authors interpreted this alternating pattern as a control impairment of the lexico-semantic system which involved an inhibition of one of the languages. Other clinical observations showed paradoxical recovery with specific use of L1 or L2 alternatively over time [100].

The neural mechanisms involved in the language selection and control processes are probably subserved by a non-linguistic fronto-subcortical network. Evidence for this was found in some clinical studies [21, 23, 44, 85, 101, 102].
Functional imaging studies manipulating language control components (language selection/switching, interpretation/translation, interference, etc.) reported activation in the basal ganglia and particularly in the left caudate nucleus (CN) [86, 103–108]. Activation was also observed in the anterior cingulate cortex (ACC) [86, 103, 107, 109], the prefrontal areas (mainly the dorsolateral prefrontal cortex, DLPFC), the inferior frontal gyrus (IFG) [51, 52, 86, 109–114] and the supramarginal gyrus [51, 107, 115]. The fact that a large number of brain regions might participate in language control confirms to some extent prior clinical observations and suggests that the system depends more on a widely distributed system rather than on a circumscribed one. However, this large number of areas can also be explained by the relatively large number of differing paradigms that reflected to a greater or lesser extent language selection processes.

Currently, the existence of a brain network dedicated specifically to language control is still debated in view of the fact that all these regions participate in other linguistic and nonlinguistic cognitive processes. To give some examples only, the basal ganglia (including the CN) participate in planning and execution of plans to achieve goals [116], acquisition of orthographic representation [117], learning and reinforcement of a stimulus-associated response [118–120]. Other studies showed that the activation in the CN correlated with the level of cognitive difficulty in executive function tasks [121, 122]. The ACC is known to have a major involvement in conflict detection/monitoring, response inhibition [123–126] and error detection [127, 128]. Others suggest that the ACC is involved in conflict detection and the DLPFC is involved in conflict resolution [129–131]. The involvement of the prefrontal cortex (PFC) in divers language tasks (e.g. semantic categorisation, word generation and recognition or rhyme detection) [132–135], working memory and attention [136–138], and cognitive control tasks [139–141] has repeatedly been reported. Finally, activity of the supramarginal gyrus has also been found in various linguistic contexts including phonological processing [142–146], vocabulary acquisition [147] and language switching [107, 148–150]. Accordingly, without any claim of exhaustiveness in this short overview, it appears that brain areas found by functional imaging techniques during para-

digms manipulating language selection/control show a significant overlap with areas involved in general cognitive control and attention mechanisms. For this reason, it can be suggested that the language selection processes are strongly dependent on the cognitive control networks (for review see [151]).

How specific is the language selection network?

The question remains as to whether or not some of the areas found outside the classical language modules are more specifically engaged during language selection processes and, if yes, when and how they are involved. Recently, a model based on the language-switching literature suggested that an executive network (including the PFC, ACC, CN) controlled a language subcircuit that included the postero-temporal areas, supramarginal and angular gyri, Broca’s area and the superior longitudinal fasciculus [152]. Also, in previous electrophysiological research by Khateb et al., the comparison between picture naming in a language-selection bilingual context and a lexical monolingual task-selection context showed differences in brain electrical responses around 220 to 300 ms after the critical stimulus. The origin of this difference was found to be in left hemisphere areas (left middle frontal-precentral, supramarginal and angular gyri) suggesting that differences between language and task selection rely, at least partially, on left hemisphere language areas [153]. Later, Abutalebi et al. used fMRI data from a similar paradigm and highlighted differences in the left CN, the ACC, part of the PFC and the IFG [86]. However, these data left some doubts as to whether this network was specifically involved in language selection, because the control selection task also necessitated lexical processes (the subject had to name either the stimulus or the related verb, depending on the cuing). More recently, a different control-selection task (judging letters or numbers) confirmed in event-related potential (ERP) data the presence of some differences between language control network and more general cognitive control network [154].

In conclusion, some data suggest that several subcortical and cortical regions including the CN, the ACC, and the PFC belong to a network that takes part in the selection of a target language. However, a clear neural model of language control cannot be definitely established at this stage of our knowledge. It appears from research in bilinguals that there are clearly some structures that are engaged during the selection of both L2 and L1, and others that are recruited only during the use or the processing of the weaker non-dominant L2. Hence, we will be tempted to propose that language selection gradually uses specific brain modules of the general cognitive control system as a function of the necessary cognitive demands that are not only determined by the task difficulty but also by the level of proficiency and automaticity in L2.

References


