
The cognitive aspect of second language learning, specifically by immersion method, is discussed from a biological perspective. The approach taken is that of "connectionism," a recently-developed theoretical and experimental approach to human cognition. It is argued that while general cognitive functioning is unaffected by language immersion, immersion students are subject to shifts in information processing strategy toward the visual system and away from the auditory system, due to low speech comprehension ability, and that development of such a bias is a natural outcome of brain structure and the mechanics of information processing in the human brain. It is further proposed that any cognitive advantages and disadvantages accrued by learners in immersion situations are determined by the information processing demands of the particular measures of cognitive performance used. Academic achievement results for this group are also explained by how achievement is defined and success measured, from the perspective of information processing. Contains 24 references. (MSE)
DOES LEARNING A SECOND LANGUAGE UNDER IMMERSION CONDITIONS MANDATE A SHIFT TOWARDS VISUAL PROCESSING OF INFORMATION?

Anne Hawson

Centre for the Study of Curriculum and Instruction
University of British Columbia
Vancouver, Canada


Note: This is a DRAFT copy only
Introduction

Becoming bilingual has long been reputed to bring about alterations in general cognitive functioning, but there has been much debate as to whether these alterations are positive or negative in consequence (see Hawson, 1996, in press-a; and Reynolds, 1991; for recent summaries and reviews of the literature). Reynolds (1991) has even questioned whether observed differences in cognitive performance are real (p. 159). The intent of this paper is to contribute to this ongoing debate by examining the issues from a biological perspective so that a more detailed picture of what is happening to cognitive functioning under immersion conditions can be built up. The particular biologically based approach which will be employed is connectionism, to be understood in its computational neuroscientific sense. From this connectionist standpoint, it will be argued (as it has been in Hawson, 1996, in press-a, in press-b) that general cognitive functioning as such is unaffected by language immersion, but that immersion students are subject to shifts in information processing strategy brought about by their language learning circumstances. Specifically, it will be contended that immersion in a second language will promote a shift in information processing strategy towards the visual system and away from the auditory system due to low speech comprehension ability, and that the development of such a bias is the natural outcome of

1 Though "connectionism" has taken on other meanings in the past decade or so (see Churchland & Sejnowski, 1992, p. 6), the sense in which it is to be understood here is that defined by Paul Churchland (1989): "a recently developed theoretical and experimental approach to the phenomena of human cognition that is at once (a) naturalistic, (b) reductionistic, and (c) capable of explaining both the radical plasticity of human consciousness, and its intricate dependence on the extended cultural surround" (p. 130). In connectionist theorizing, "plasticity" is roughly equivalent to ability to learn. Learning itself is considered to be the process of the long-lasting changing of synaptic weights, as well as the outcome of these changes (see Hawson, in press-a for a synopsis of the connectionist view of learning, and Churchland, 1989, chaps. 9 and 10. and Churchland & Sejnowski, chap. 5 for more detailed information).
brain structure and the mechanics of information processing in the human brain. It will also be argued that any "cognitive" advantages and disadvantages accrued by learners in immersion situations are determined by the information processing demands of the particular measures of cognitive performance employed. Academic achievement results for this group will similarly be explained through information processing considerations of how such achievement is defined by measures of academic success.

The Case for a Shift in Information Processing Strategy towards the Visual System

Multimodality of speech comprehension
In recent years, many of the anatomical links within and between various systems in the brain have been mapped, and it has been shown that what have been in the past considered unitary systems are in fact cooperative systems of systems operating at different levels of spatial dimension. Indeed, Antonio Damasio (1994) has described the brain itself as a "supersystem of systems. Each system is composed of an elaborate interconnection of small but macroscopic cortical regions and subcortical nuclei, which are made up of microscopic local circuits^2, which are made up of neurons, all of which are connected by synapses" (p. 30). For an example of a sensory system which has been extensively investigated and the integrated nature of whose inner workings is now to a large extent understood, see Zeki's (1992) explanation of the visual system of the human brain.

The knowledge system which subserves language is no exception to this subdivided but integrated pattern. (See Kosslyn & Koenig, 2

^2 Here Damasio is distinguishing between the terms "system" and "circuit" (or "network") which are often used interchangeably in the literature. "Circuits" he considers to denote microscopic systems, while "systems" are considered macroscopic within the context of brain research.
1992, p. 249, for an elaborate diagram of the subsystems identified within the language system and their interconnections.) As Jacobs (1988) has pointed out, language is dependent not only on various subsystems within a particular sensory modality, but on various sensory modalities:

Language depends on linguistic information and, as a multimodal sensory system, is not particular about the mode of input it receives; it will make use of what is available. In the absence of one mode (e.g. audition), the other (e.g. vision) will supply the necessary linguistic information. Language is subserved by the senses insofar as it depends on the linguistic information they provide, but it also transcends them because it is not modality specific. (p. 312)

This multimodal character of language has been demonstrated empirically by language research. It has been shown, for example, that the vibratory and tactual information of the neck and face during articulation, as well as visual information of the face and mouth can communicate language (see Jenkins, 1991). Indeed, visual information appears to be particularly important to speech perception, an aspect of language which "is typically ignored by generative linguists (phoneticians excluded)" (Jacobs, 1988), but which it is of crucial importance for learners in language immersion situations to master if fluency of comprehension and production are sought.

An instance of the importance of visual information to speech perception can be found in the results of a study by Danhauer, Crawford, and Edgerton (1984). These researchers discovered that there was an improvement in performance scores on an English-based

---

3 Throughout this paper, the position is taken that speech perception is fundamental to true speech comprehension, though it is recognized that the latter is aided by the contextual information available in a communicative situation.
nonsense syllable test for all groups (English, Spanish/English bilinguals, and predominantly Spanish speakers) when the test was administered in a face-to-face presentation rather than a controlled auditory presentation only. Furthermore, in cross-language studies on adults, it was discovered that visual information not only affects speech perception, but can actually override auditory information in situations in which the visual and auditory information do not correspond (see Werker, 1991). Hence, though audition would be expected to be the dominant modality in speech perception, vision evidently plays an important role.

Albeit multimodal, there is no doubt that the auditory and visual sensory systems are by far the most important to speech perception and therefore to speech learning, and, as Jacobs has noted, when one mode of input is unavailable, the other will supply whatever linguistic information it can. Thus the contributions of visual and auditory information can be conceptualized as varying along a continuum between fully auditory and fully visual, with the particular position on the continuum at any time being dictated to a large extent by environmental circumstances. Given this importance of visual information to speech perception, it is of interest to inquire into the kind of information vision can give on speech.

According to Alvin Liberman, what is supplied by both the visual and auditory systems with respect to speech is information on phonetic gestures, "organized movements of one or more vocal-tract structures that realize phonetic dimensions of an utterance" (see Fowler & Rosenblum, 1991, p. 34). Liberman and Mattingly (1985), however, have noted that these gestures "must be seen, not as peripheral movements, but as the more remote structures that

It is to be expected that individual characteristics such as modality preference would interact with environmental circumstances in determining how much the auditory and visual systems of the person contribute.
control the movements. These structures correspond to the speaker's intentions" (p. 23). What this means is that, visual and auditory information may be used interchangeably by the speech perception system of the brain because they both provide data on the speaker's intentions with respect to the speech sounds being produced. Furthermore, in the human brain, it would appear that the visual and auditory systems have been anatomically set up to coordinate their input; Kosslyn and Koenig (1992) have pointed out that the structure upon which the allocation of auditory attention depends, the inferior colliculus, projects to motor neurons in the superior colliculus, a visual reflex centre, so that auditory information can produce a visual fixation of the source of the sounds. Thus, "one typically pays attention to a single object, registering its appearance and sounds at the same time" (p. 217). In the case of someone speaking a language, then, visual and auditory data related to the utterance are attended to and stored simultaneously, at some point in the process converging on the microsystem which stores data on phonetic gestures. The main point being made here is that, under normal circumstances, the auditory and visual systems work in concert in locating and perceiving sounds, including speech sounds, and, in consideration of this fact, it seems reasonable to propose, as Jacobs has done, that, when information is deficient in some way (low intensity, or garbled, or even not previously encountered, as might be the case for second language learners) corroborating

5 The description of phonetic gestures here has resonance with Damasio's (1992) definition of "dispositional representations [which] exist in potential state, subject to activation, like the town of Brigadoon":

What I am calling a dispositional representation is a dormant firing potentiality which comes to life when neurons fire, with a particular pattern, at certain rates, for a certain amount of time, and toward a particular target which happens to be another ensemble of neurons.... Innate knowledge is based on dispositional representations in hypothalamus, brain stem, and limbic system.... Acquired knowledge is based on dispositional representations in higher-order cortices and throughout many gray-matter nuclei beneath the level of the cortex. (pp. 103-105)
visual information would be sought out to aid perception of the speech sound. Similarly, a higher intensity acoustic signal might be sought out if visual information with respect to an utterance is unavailable.

What is being suggested by the research above is that, in situations where the auditory system is not supplying enough information for understanding (and the appropriate decision-making that comes through adequate understanding), more attention will be allocated to the visual system. When this happens, cellular activity in the visual cortex will be enhanced thereby enhancing the possibility of learning through this system, while activity in the auditory cortex may be suppressed (Haxby et al., 1994) with a commensurate reduction in learning ability via the auditory system (see Hawson, 1996, in press-a, for more details on the effects of attention on learning).

The usefulness of a preferential allocation of attention to the visual system for speech comprehension extends well beyond that of enhancing speech perception, however. Quite apart from supplying linguistic information on phonetic gestures, visual information would be of value to a second language learner in an immersion situation by virtue of the fact that it could provide semantic information through observation of the speaker's mood, for example, or of body gestures. In other words, the visual system could provide linguistic context which would aid comprehension of any utterance. Moreover, as has been discussed previously (Hawson, 1996, in press-a, for more details on the effects of attention on learning).

6 The definition of attention which is being used here is the following: A facilitatory mechanism for the selection of information for further processing. It works in cooperation with expectation and intention to act, and can be driven by voluntary-or stimulus-initiated mechanisms (see Hawson, 1996, for an in-depth discussion of the nature of attention; also Hawson, in press-b, for a less technical version). The allocation of attention to an area of the brain has been shown to enhance cellular activity and hence learning in that area.
1996, in press-a), it is likely that the semantic content in the visual information would be more familiar to the learner than that of the auditory system. Familiarity of semantic content would allow the information to be processed via the fast information processing route that Raichle et al. (1994) have recently identified through PET (Positron Emission Tomography) scan studies on the human brain. The familiar visual input would also draw attention to the visual system in part because it is processed faster, but also because, being familiar, it has already acquired value for the learner. It is recognized in attention research circles that what is allocated attention in input and in what order is determined by considerations of what has value (or salience) to the perceiver. This value or salience can be either innately specified or learned (see Hawson, 1996; Edelman, 1992), and, as far as language learning is concerned, the determination of which speech sounds have value starts young.

Organizing and reorganizing the speech perception system
It has been demonstrated that, even in early infancy, humans can distinguish native and non-native phonemes as categories. Hence, it would appear that experience is not necessary for their perception (Kosslyn & Koenig, 1992, p. 221; Werker & Tees, 1992). What experience with a first language does, therefore, is to enhance the perception of phonemes and complex groupings of phonemes (as well as other phonological patterns) which are of value in the learner's environment. If a particular phoneme is not used in the language a child is exposed to, it later becomes difficult to perceive. This

Raichle et al. have identified two mutually exclusive information processing routes in the brain, one which is active during the processing of novel/naive tasks and is slow, and another which comes into play during the processing of practised tasks and is fast. Though it would appear from the evidence that the two systems do not work simultaneously, under normal processing conditions, there would be a continual interplay between the two systems. (See Hawson, 1996, and in press-a for further information on these information processing routes.)
"weeding out" of valueless phonemes and the development of sensitivity towards the speech sounds of the native language takes place within the first year of life (Werker & Tees, 1984, 1992; Werker, 1991). Looking at this issue of value from an interesting perspective, Kosslyn and Koenig (1992) comment that "[a]pparently, our auditory systems develop certain insensitivities (which lead us not to make discriminations; i.e. to group within a category) because this is useful" (p. 222). This results in "overall, a sharpening of distinctions that count, and a grouping together or collapsing of sounds that are not important in the child's environment" (Jenkins, 1991, p. 436). The end product of linguistic experience with the mother tongue is that certain combinations of sounds are stored as having high value, and when activated auditorily, these stored representations (or mini-systems) can subsequently be perceived preferentially.

This acquired specificity of the auditory system for the speech sounds of the first language constitutes a significant perceptual problem for the second language learner. In the beginning stages of exposure to the second language, sound categories particular to that language, including previously suppressed phonemes, will not be easily perceived by the learner because, never having been present in the environment, they have never been ascribed value. As a consequence, they will not be accorded attention and hence will not receive the facilitatory information processing boost that familiar, valued sound groupings receive. According to Kosslyn and Koenig (1992, pp. 219-223), it is because of the lack of facilitation of the processing of sound groupings prevalent in a target language that second language learners have difficulty hearing the speech sounds of the unfamiliar language. In other words, these learners can be considered to have temporary phonological deficits. Such difficulties would explain the results of the previously mentioned study by Danhauer, Crawford and Edgerton (1984), the main findings of which were that predominantly Spanish speaking subjects had greater difficulty perceiving the
speech sounds on an English-based nonsense syllable test than did either English or fluent English/Spanish bilingual speakers (whose performances were nearly equivalent) thereby rendering them likely to be judged hard of hearing if tested using this type of instrument. As noted above, this was especially so when the test was presented under controlled auditory only conditions. Significantly for the premises of this paper, the improvement in score for predominantly Spanish speakers seen for face-to-face presentation (i.e. with visual information of mouth and face motions available) was more than double that of the other two groups (14% versus 6%). Given the multimodality of speech perception, it is highly likely that a deficient performance of the auditory speech processing system of this sort will promote a strategic shift in allocation of attention to the visual system in second language learners in immersion situations.

Some connectionist modelling insights

Although the apparently natural tendency of the brain to allocate attention to the visual system would appear to be counter productive to the task at hand, that of reorganizing (or reeducating) the auditory speech processing system to value and thus facilitate the processing of previously suppressed or unknown speech sound categories, there is a possibility that a temporary withdrawal of voluntary attention from the auditory system would be beneficial. In order to understand why, it is necessary to consider the connectionist concepts of "supervised" and "unsupervised" learning, concepts which were developed to enable machine models to mimic certain aspects of biological learning. In supervised learning, the model is set up such that a network or learning system has an source of error evaluation which monitors the difference between the actual output of the net and the desired output, and the changes in the connection strengths (model "synapses") are altered so as to bring the performance of the model progressively closer to the desired output. In other words, there is an information source or "teacher" external to the system which
guides the learning process. It has been argued in Hawson (1996) that the action of this information source is analogous to the priming action of expectation and intention to act, both external sources of information rooted in past experience, on attention. Thus, in voluntary attention allocation in which expectation and intention to act are actively engaged, what is given the facilitatory processing boost of attention, and therefore what is learned more easily, is determined to a large extent by past experience. But in the case of a learner trying to encode new sound groupings, the selective action of expectation and intention to act would be ill-informed to facilitate the processing of the new sounds due to the fact that it is based on information valued in first language processing. That is to say, in the initial stages of second language learning, voluntarily allocating attention to the speech processing system would prime and therefore promote the processing of speech sounds particular to the first language, and would therefore work against the encoding of new speech sounds.

Unsupervised learning in machine models, on the other hand, is learning that occurs without recourse to an external evaluation of the output, though some unsupervised nets can monitor their own performance through internal feedback. The usefulness of unsupervised nets resides in this ability to learn without instruction as to what is valuable for categorization in the input. As Churchland and Sejnowski (1992) put it "the net structures itself to represent whatever systematicity it can find in the input" (p. 98). This kind of net can be employed as a feature detector in a machine "whose sensory input must be encoded in some perspicuous fashion before it is sent on for use in such tasks as pattern recognition and motor control" (ibid.). In line with the argument laid out above with respect to the involvement of attention in supervised learning, it could be said that, in an unsupervised learning situation, which connectionist modellers consider to be more biologically plausible than supervised learning, learning is not driven by top-down, voluntary attentional
mechanisms. Whatever attention is accorded to the input would be more stimulus-initiated, and would therefore be geared to the discovering of "systematicity" or regularities in the speech input, features which should be determined largely by frequency of occurrence. This emphasis on discovery rather than guidance in unsupervised learning, which it is argued here and elsewhere (Hawson, 1996) is equivalent to learning without focused, voluntary attention, would make it a more suitable learning mode in the initial stages of encoding the sound groupings or phonological patterns of a second language since the value of previously unencountered sound groupings would be built up gradually with exposure to the language. Of course, external indicators of value such as emphasis on the part of the speaker would always be interactive with the stimulus-initiated categorization. Hence, a purely unsupervised mode of learning would almost certainly never be fully operative. Once phonological awareness of the second language has been built up, speech perception and therefore comprehension should be considerably enhanced, and dependence on visual processing of information should diminish accordingly.

The Shift in Information Processing Strategy Hypothesis
On consideration of the multimodality of speech perception, the processing of familiar information, and the gathering contextual semantic clues to speech, as well as insights from the connectionist modelling of learning outlined above, it has been proposed (Hawson, 1996, in press-a, in press-b) that learners in the initial stages of second language immersion undergo a shift in information processing strategy towards the visual system and away from the auditory. It is also proposed that this shift is largely

8 In earlier papers, Hawson in press-a, in press-b, this hypothesis has been termed simply the "shift in information processing" hypothesis.

9 The formal statement of the shift in information processing hypothesis reads as follows: It is hypothesized on the basis of a connectionist perspective on learning that, during the initial
involuntary, initiated, and perhaps even mandated, by the brain itself as a consequence of its normal manner of functioning under immersion language learning circumstances. As would be expected of a dynamical "supersystem of systems," any deficiency in the information gathering capability of the auditory system would automatically trigger a compensatory increase in the contributions of other interacting systems whose informational contribution would be relevant to the problem or task at hand. It is to be expected, then, that the other major contributing system to speech perception and comprehension, the visual system, would exhibit an enhancement of its activity due to an increase in allocation of attention. This enhancement of activity would mean that information processing tasks that are readily accomplished by means of visual analysis would be performed more easily than normal (for each individual learner), while tasks requiring auditory processing may be more difficult.

Evidence to support a facilitation of visual processing can be found in a study reporting high mathematics achievement by second language learners, especially in the early stages of residence in a new country (see Collier, 1987), mathematics (particularly computational mathematics) being one of the few academic subjects in which language content is minimal, and in which problems are analyzable by visual means (see Hawson, in press-a, in press-b, 1996). Furthermore, it has been argued (ibid.) that psychological testing of cognitive ability provides additional support for the notion that, especially in the beginning stages of second language learning, the visual processing ability of second language learners is functioning adequately. Most intelligence testing items given to stages of learning in immersion situations, learners will adopt, whether consciously or subconsciously, a shift in information processing strategy such that visual processing of information will be favoured over auditory processing, and that this shift will bring about an increased ability to learn via the visual system while it is in effect. (Hawson, 1996)
second language learners today are heavily dependent on visual analysis in order to rule out the biasing of results due to language deficiencies, and, based on such testing, Diaz (1985) and Hakuta (1987) have shown the "cognitive" ability of students at the low end of the bilingual ability scale to increase as a function of increasing knowledge of the second language. These results, added to the fact that mathematics has always been considered a high cognitive content academic subject, should nullify any suggestion that the general cognitive ability of second language learners in immersion situations is compromised at any stage in the learning of the second language. Indeed, given the dependence mentioned above of studies reporting positive cognitive results on testing dependent on visual analysis of tasks, these learners cannot be considered to be in either a general cognitive deficit or benefit position as has been postulated by different bodies of second language research (see Reynolds, 1991) whatever their level of bilingual ability. That is to say, the "benefits" or "deficits" accorded to them are task related. While they may experience some difficulties accomplishing tasks dependent on auditory processing, it may be easier than normal (for the individual student) to perform tasks requiring visual analysis. These effects should be especially prominent during the early stages of second language learning and should diminish as auditory comprehension of the language increases. Whether information processing tendencies will return to pre-second language learning states is debatable, however. Because of the incredible plasticity of the brain, it seems probable that some effects of the strategic shift towards the visual would remain in the form of an increased tendency and ability to process information visually.

**Educational implications**

The most important implication of the shift in information processing strategy hypothesis is that the general cognitive functioning of second language learners, at whatever stage of bilingual ability they have reached, should not be a source of
concern for educators. But good overall cognitive functioning may not equate to good academic achievement. Good academic achievement is closely tied to good reading ability, which is dependent not only on vision but also on audition. It has been shown by many studies that reading disabilities in monolinguals are closely associated with problems of phonological awareness, and can be aided by raising such awareness (see McShane, 1991, chap. 8). Indeed, some recent studies in cognitive neuroscience have linked difficulties in encoding certain temporal aspects of speech with difficulties in reading (see Miller, Delaney, & Tallal, 1995). Second language learners who develop a pronounced and persistent bias towards processing information via the visual system may be doing so because of similar auditory perception difficulties, and may consequently be in danger of developing reading difficulties. That reading is not an academic subject area in which these learners generally perform well can be seen from the universally low reading achievement scores (the lowest of all academic achievement measures listed, and well below school system norms) across all grades of ESL students in Collier's (1987) study, regardless of length of residence in the country. Of course, once reading ability is compromised, academic achievement is undoubtedly endangered. Hence, testing the second language learner's phonological awareness of the second language at regular intervals and initiating remedial interventions if necessary may be beneficial in warding off reading-based problems in academic achievement for these students.

References


10 This argument has been developed more fully in Hawson 1996, in press-a, in press-b, where minority group second language learners whose first language is not prevalent within a society are identified as being especially at risk.


