Neurolinguists, Beware! The Bilingual Is Not Two Monolinguals in One Person

FRANÇOIS GROSJEAN

Université de Neuchâtel, Switzerland

Two views of bilingualism are presented—the monolingual or fractional view which holds that the bilingual is (or should be) two monolinguals in one person, and the bilingual or wholistic view which states that the coexistence of two languages in the bilingual has produced a unique and specific speaker-hearer. These views affect how we compare monolinguals and bilinguals, study language learning and language forgetting, and examine the speech modes—monolingual and bilingual—that characterize the bilingual’s everyday interactions. The implications of the wholistic view on the neurolinguistics of bilingualism, and in particular bilingual aphasia, are discussed.

This paper is divided into three parts. In the first, a particular view of bilingualism, termed the monolingual view, is discussed and criticized. It holds that bilinguals are (or should be) two monolinguals in one person and that they can therefore be studied like any other monolingual. In the second part, a different, more recent, theory is examined. The bilingual (or wholistic) view holds that the bilingual is NOT the sum of two complete or incomplete monolinguals; rather, he or she has a unique and specific linguistic configuration. This view is described and discussed with reference to three different domains: the comparison of monolinguals and bilinguals, language learning and forgetting, and the speech modes—monolingual and bilingual—that bilinguals find themselves in during their everyday interactions. In the third part, the implications of this view on the neurolinguistic study of bilingualism are discussed. The assessment of the linguistic and communicative abilities of bilingual aphasics before and after injury is examined, and suggestions for the examination of patients are proposed.

This paper was written while the author was on leave at the Université de Neuchâtel, Switzerland. Its preparation was supported in part by grants from the National Science Foundation (BNS-8404565) and the Department of Health and Human Services (RR 07143). Requests for reprints should be addressed to: François Grosjean, Laboratoire de traitement du langage et de la parole, Université de Neuchâtel, 2000 Neuchâtel, Switzerland.
(1) THE BILINGUAL IS NOT TWO MONOLINGUALS IN ONE PERSON

A strong version of the monolingual (or fractional) view of bilingualism is that the bilingual has (or should have) two separate and isolable language competencies; these competencies are (or should be) similar to those of the two corresponding monolinguals; therefore, the bilingual is (or should be) two monolinguals in one person. This view, which is prevalent among many researchers, educators, and bilinguals themselves, is a result of the strong monolingual bias that has been prevalent in the language sciences. Monolinguals have been the models of the "normal" speaker-hearer, and the methods of investigation developed to study monolingual speech and language have been used with little, if any, modification to study bilinguals.

This monolingual view of bilingualism has had a number of negative consequences (Grosjean, 1985a). The first is that bilinguals (which we define as those people who use two or more languages in their everyday lives) have usually been described and evaluated in terms of the fluency and balance they have in their two languages. The "real" bilingual is seen as the person who is equally and fully fluent in two languages; he or she is the "ideal," the "true," the "balanced," the "perfect" bilingual (see Bloomfield, 1933; Thiery, 1978). All the others, who in fact represent the vast majority of people who use two languages in their everyday lives, are "not really" bilingual or are "special types" of bilinguals; hence the numerous qualifiers found in the literature: "dominant," "unbalanced," "semilingual," "alingual," etc. This search for the "true" bilingual has used traditional language tests as well as psycholinguistic tests which are constructed around the notion of "balance"; invariably the "ideal" bilinguals are the ones who do as well in one language as in the other. All others are somehow "less bilingual" and are put into an indeterminate category.

A second consequence of the monolingual view is that language skills in bilinguals have almost always been appraised in terms of monolingual standards. The tests used with bilinguals are often quite simply the tests employed with the monolinguals of the two corresponding language groups. These tests rarely take into account the bilingual's DIFFERENTIAL NEEDS for the two languages or the DIFFERENT SOCIAL FUNCTIONS of these languages (what a language is used for, with whom, where, etc.; see Fishman, 1965). Many monolingual tests are quite inappropriate to evaluate the language skills of bilinguals; others need to be adapted substantially.

A third effect of the monolingual view is that the cognitive and developmental consequences of bilingualism have received close scrutiny. Because this view considers bilingualism as the exception, when, in fact, half of the world's population is bilingual, it has long been held that the
knowledge and use of two languages will have profound (negative or positive) effects on a person's psychology and cognitive functioning. And yet, despite innumerable studies, the "effects" literature has never been able to factor out the sole bilingualism variable from a host of other linguistic and sociocultural factors: the language used to test the bilinguals; the use they make of their languages; their socioeconomic background; the monolingual control group, etc. Studies simply have not been able to show a direct, unambiguous, causal relationship between using two (or more) languages in one's everyday life and various cognitive or developmental effects.

A fourth consequence of the monolingual view is that the contact of the bilingual's two languages is seen as accidental and anomalous. Because bilinguals are (or should be) two separate monolinguals in one person, covert or overt contact between their two languages should be rare. The two language systems should be autonomous and should remain so at all times. If there is contact, it is accidental and simply the result of language interference; borrowing and code-switching, which occur in conversations between bilinguals, are either included in the interference category or are explained away as the product of careless language.

A fifth consequence is that research on bilingualism is, in large part, conducted in terms of the bilingual's individual and separate languages. For example, psycholinguists have been interested in how the bilingual's two languages are activated one at a time, and hence have paid little attention to the simultaneous activation of the two languages as in the case of borrowing and code-switching. Linguists have shown little interest in the bilingual's language competence in the Chomskyan sense, maybe because the bilingual can never be an "ideal speaker-hearer" in the same way that the monolingual supposedly can; there is no real acceptance among linguists that the bilingual's two grammars can differ from the corresponding monolingual grammars or that language competence (and especially first language competence) can actually change when another language is acquired and begins to dominate. Finally, many neurolinguists and speech therapists are still using standard monolingual tests with their bilingual subjects; these tests rarely take into account the situations and domains the languages are used in, the skills covered by these languages, or the amount and type of code-mixing normally produced by the bilingual. Unfortunately, much of what we know about bilingualism today is tainted—in part at least—by a monolingual, fractional, view of the topic.

A final effect is that bilinguals rarely evaluate their language competencies as adequate. They often assume and amplify the monolingual view and hence criticize their own language competence: how many times have bilinguals reported that they neither speak nor write their different languages adequately! Other bilinguals strive to reach monolingual norms and still others hide their knowledge of their "weaker" language.
To conclude this section, we should stress how important it is to have a clear understanding of the monolingual view of bilingualism and of the impact it has had on our thinking. We may then be ready to consider bilinguals, not as two monolinguals in one person, but as different, perfectly competent speaker-hearers in their own right.

(2) THE BILINGUAL AS A COMPETENT BUT SPECIFIC SPEAKER–HEARER

A bilingual (or wholistic) view of bilingualism proposes that the bilingual is an integrated whole which cannot easily be decomposed into two separate parts. The bilingual is NOT the sum of two complete or incomplete monolinguals; rather, he or she has a unique and specific linguistic configuration. The coexistence and constant interaction of the two languages in the bilingual has produced a different but complete linguistic entity. An analogy comes from the domain of track and field. The high hurdler blends two types of competencies, that of high jumping and that of sprinting. When compared individually with the sprinter or the high jumper, the hurdler meets neither level of competence, and yet when taken as a whole the hurdler is an athlete in his or her own right. No expert in track and field would ever compare a high hurdler to a sprinter or to a high jumper, even though the former blends certain characteristics of the latter two. A high hurdler is an integrated whole, a unique and specific athlete; he or she can attain the highest levels of world competition in the same way that the sprinter and the high jumper can.

In many ways, the bilingual is like the high hurdler: an integrated whole, a unique and specific speaker–hearer, and not the sum of two monolinguals. He or she has developed competencies (in the two languages and possibly in a third system that is a combination of the first two) to the extent required by his or her needs and those of the environment. The bilingual uses the two languages—separately or together—for different purposes, in different domains of life, with different people. Because the needs and uses of the two languages are usually quite different, the bilingual is rarely equally or completely fluent in the two languages. Levels of fluency in a language will depend on the need for that language and will be domain specific (hence the "fossilized" competencies of many bilinguals in each of their two languages).

Because the bilingual is a human communicator, he or she has developed a communicative competence that is sufficient for everyday life. This competence will make use of one language, of the other language, or of the two together (in the form of mixed speech) depending on the situation, the topic, the interlocutor, etc. The bilingual’s communicative competence cannot be evaluated through only one language; it must be studied instead through the bilingual’s total language repertoire as it is used in his or her everyday life.
A number of areas of study are affected by this view of bilingualism. Each will be discussed below.

(a) Comparing Monolinguals and Bilinguals

A wholistic view of bilingualism should lead to a more precise and fairer comparison of bilinguals and monolinguals. The comparison will stress the many specificities of the bilingual:

—the structure and organization of the bilingual’s language competencies; it may well be that these competencies are in some ways different from those of the two corresponding monolinguals;

—the structure and organization of the bilingual’s mixed language competence, that is, the language systems that are activated when the bilingual is in a bilingual (mixed) speech mode and is borrowing and code-switching with other bilinguals;

—the bilingual’s language processing systems when the language input and output are monolingual (as when the bilingual is speaking to monolinguals; we know that in such cases the other language is never totally deactivated);

—the linguistic and psycholinguistic operations involved in producing and perceiving mixed speech.

But the comparison of bilinguals and monolinguals will also need to stress the many similarities that exist between the two at the level of communicative competence. We hypothesize that the STABLE BILINGUAL (as opposed to the person in the process of acquiring or restructuring a language) has developed a communicative competence that is equivalent to that of other speaker-hearers, be they monolingual, bilingual, or multilingual. This is the case even though the outward manifestations of this competence may at first appear quite abnormal to the monolingual researcher, as in the case of mixed speech, which is so often seen as a reflection of semilingualism, alingualism, or even language disorder! To confirm this hypothesis, we will need to develop new testing procedures; traditional language tests that put more stress on the FORM of the language than on the speaker’s ability to communicate in context are not appropriate. We will also need to study in more detail how monolinguals and bilinguals implement their communicative competence: the former with just one language, and the latter with two (or more) languages, used separately or together, depending on the speech mode they are in (see below).

(b) The Wax and Wane of the Bilingual’s Languages

A second hypothesis can be proposed: a person can go in and out of bilingualism, can shift totally from one language to the other (in the sense
of acquiring one language and forgetting the other totally), but will never depart from a necessary level of communicative competence needed by the environment, except in transitional periods of language learning or language restructuring. Because bilinguals, like monolinguals, have an innate capacity for language and are by essence communicators, they will develop a formal competence in each of their languages to the extent needed by the environment (the competence in one language may therefore be quite rudimentary, as the interlanguage literature has shown (Corder, 1967; Selinker, 1972)). However, they will always maintain a necessary level of communicative competence: new situations, new environments, new interlocutors will involve new linguistic needs in one language, in the other, or in both simultaneously, and will therefore change the language configuration of the person involved, but this will in no way modify his or her communicative competence. After a period of readjustment (or language restructuring) the person will meet his or her new communicative needs to the fullest.

It is critical to differentiate between the process of restructuring a language and the outcome of restructuring, in other words, between becoming bilingual or readjusting one’s bilingualism and attaining stability in one’s bilingualism. It is also important to study what happens to the two languages (and to the interaction of the two) during this period of readjustment. In the long run, the really interesting question is how the human communicator adjusts to, and uses, one, two, or more languages—separately or together—to maintain a necessary level of communicative competence. Of much less interest is the level of formal competence reached in each language when taken individually and out of context. Unfortunately, too much stress has been put on the latter in bilingual research.

(c) The Bilingual’s Speech Modes

In their everyday lives, bilinguals find themselves at various points along a situational continuum which induce a particular speech mode. At one end of the continuum, bilinguals are in a totally monolingual speech mode: they are speaking to monolingual speakers of either language A or of language B and therefore have to restrict themselves to just one language (A or B). At the other end of the continuum, they are with bilinguals who share their two languages (A and B) and with whom they normally mix languages (code-switch and borrow): they are here in a bilingual speech mode. For convenience, we will refer to the two endpoints of the continuum when speaking of the monolingual or bilingual speech modes, but we should keep in mind that intermediary modes exist between the two.

Before describing the endpoints we should note two things. First, bilinguals differ among themselves as to the extent they travel along the
BILINGUALS: NOT TWO MONOLINGUALS

continuum; some rarely find themselves at the bilingual end (purists, language teachers, etc.) whereas others rarely leave this end (bilinguals who live in tight-knit bilingual communities where the language norm is mixed language). Second, it is critical to know which speech mode a bilingual is in before making any claims about the individual’s language processing or language competence. For example, what might be seen as the accidental (or permanent) interference of one language on the other during language production, may in fact be an instance of borrowing or code-switching in the bilingual speech mode. Rare are the studies that clearly indicate the speech mode the bilinguals were in when they were recorded or tested; as a consequence, many unfounded claims have been made about the bilingual’s languages and speech.

In the monolingual speech mode, bilinguals adopt the language of the monolingual interlocutor and deactivate, as best they can, the other language. As is well known, deactivation is rarely total (Blair & Harris, 1983; Obler & Albert, 1978; Paradis, 1987), and this is clearly seen in the interferences bilinguals produce. Interferences are those deviations from the language being spoken (the base language) due to the involuntary influence of the other “deactivated” language. They can occur at all levels of language (phonological, lexical, syntactic, semantic, etc.) and in all modalities (spoken or written language). Interferences can be of two kinds: static interferences which reflect permanent traces of one language on the other (such as a “foreign accent”), and dynamic interferences, which are the ephemeral and accidental intrusions of the other language (as in the case of the accidental slip on the stress pattern of a word due to the stress rule of the other language, or the momentary use of a syntactic structure taken from the language not being spoken). These latter interferences occur more or less randomly whereas the first type are systematic.

In the bilingual speech mode, where both languages are activated, bilinguals become quite different speaker-hearers. First they “choose” a language to use with their bilingual interlocutor (we will call this the “base language”), and then they intermix the other language when needed (this is often referred to as “language mixing”). The actual choice of the base language is a function of many factors such as the participants involved, the situation, the topic, and the function of the interaction (Grosjean, 1982). Once a particular base language has been chosen, a bilingual can bring in the other language in several ways: by switching completely to that language for a word, a phrase, a sentence (this is known as code-switching) or by borrowing a word from the language and integrating it phonologically and morphologically into the base language. This type of idiosyncratic loan is called a “speech borrowing” to distinguish it from a “language borrowing” which is a word that has become part
of the base language (such as "weekend" and "pullover" in Parisian French).

Code-switching involves the complete shift to the other language for a word, a phrase, a sentence, or an utterance. French–English examples are: "On est pas assez QUICK" ("We're not quick enough"), or "J'ai l'impression d'être BACK IN THE COUNTRY" ("I have the feeling of being back in the country"). Code-switching has received considerable attention lately from researchers who have studied the psychosocial and communicative factors that underlie it as well as the grammatical constraints or rules that govern it (see, for example, Gumperz, 1982; Scotton & Ury, 1977; Pfaff, 1979; Poplack, 1980; Grosjean, 1982). It is now accepted by most that code-switching reflects linguistic and communicative strategies in bilinguals speaking to one another, that natural switches (produced in a relaxed atmosphere) are rarely marked off by prosodic markers, and that mixed discourse is understood as easily as monolingual discourse.

The other way a bilingual can mix languages is to borrow a word from the language not being spoken and to adapt it phonologically and mor-
phologically into the base language. In the sentence, "On a BRUNCHÉ chez eux" ("We brunched at their place"), the English word "brunch" is adapted phonologically and morphologically into the base language, and becomes, to all intents and purposes, a French word.

Current psycholinguistic research on code-switching and borrowing (for example, Grosjean, in press; Grosjean and Soares, 1986) is seeking answers to the following kinds of questions: How does the bilingual speaker program and execute an utterance that contains code-switches? At what point in the speech stream does the speaker switch phonetically from one language to the other? How complete is the switch? How does the bilingual listener perceive and comprehend a mixed language input? What strategies and operations lead him or her to process the utterance appropriately? How is a borrowing accessed in the appropriate mental lexicon when the acoustic–phonetic (and sometimes morphological) in-
formation signals a word from the base lexicon? etc.

Having examined the speech modes bilinguals find themselves in, it is interesting to narrow the discussion to a topic that is of interest to neurolinguists: language mixing in aphasic speech. A question which has raised some controversy (see Perecman, 1984; Grosjean, 1985b) is whether language mixing by bilingual aphasics is always a correlate of language deficit. An examination of the many case studies reported in the literature (Albert & Obler, 1978; Paradis, 1983, for example) shows that certain aspects of mixing are clearly due to deficit; among these we find the use of the wrong base language with a monolingual interlocutor (thus leading to a breakdown in communication), extensive code-switching with a monolingual (again resulting in non-communication), violating code-
switching constraints or rules, language mixing while reading a monolingual
text, failing to switch or translate upon request, etc. Rare are the case studies, however, that isolate these kinds of exceptional mixes from those that are quite acceptable in bilingual speech. To do this, one needs to know about the testing situation (who the examiners were, what languages they knew and spoke with the patient, the speech mode the patient was in during testing) and one needs information about the language knowledge and the language behaviors of the patient before injury. A question of interest concerns the kind of static, as well as dynamic, interferences which occurred in the patient’s languages before injury. It could well be that an “error” in pronunciation or the use of a wrong word in a sentence during testing simply reflected the patient’s normal interference behavior prior to injury. A second, more important question concerns the speech mode the patient was in when being tested. From the description of various case studies, one can infer that a number of polyglot aphasics were probably examined by people who knew some, if not all, of the patients’ languages. If that was the case, and the rapport between the patients and the investigators was good, then the aphasics might well have code-switched and borrowed during testing. All the more so if communication in one language proved difficult: the patient, aware of production problems in that language, might well have adopted a strategy of code-switching and simultaneous translation to enhance communication. These communicative strategies, aimed at enhancing the flow of information, make no sense with a monolingual examiner (and would therefore be true reflections of deficit), but are perfectly valid with a bilingual examiner, especially if the latter has a marked preference for the language not being used in the conversation. We can conclude then that language mixing by bilingual aphasics can have many causes. It can reflect language and conceptual deficits, but it can also be the result of communicative strategies used both by normal and impaired bilinguals. Unfortunately, most existing case studies do not allow us to disentangle these various causes.

(3) IMPLICATIONS FOR THE NEUROLINGUISTIC STUDY OF BILINGUALISM

Viewing the bilingual as a competent but specific speaker-hearer, and not as two monolinguals in one person, has some bearing on the neurolinguistic study of bilingualism, be it experimental or clinical. To illustrate this point I will discuss the assessment of bilingual aphasics, an endeavor that has been the object of considerable attention lately (see, most notably, Paradis’ Assessment of Bilingual Aphasia (1987)). The questions that will be raised, and the suggestions proposed, should extend to all domains of bilingual neurolinguistics.
(a) Describing the Bilingual Prior to Injury

Although everyone concurs that there are major difficulties involved in adequately describing a patient's bilingualism prior to injury, it is nevertheless important not to overlook certain critical questions. Some of these are:

—Which languages did the patient know before injury?
—How well did he or she know them (as a function of linguistic level, language skills, styles, etc.)?
—What were the languages used for, with whom, for what?
—What kind of interferences occurred in the patient's speech when in a monolingual speech mode? When speaking language A? Language B?
—Which of these interferences were of a static nature? Which of a dynamic kind?
—How much time did the patient spend in a monolingual as opposed to a bilingual speech mode?
—How much mixing took place in the bilingual speech mode (if and when the patient was in that mode)?
—What kind of mixing occurred: speech borrowing, code-switching, both?
—Who did the patient code-switch and borrow with?
—How good were the translation abilities of the patient? etc.

(b) Describing the Bilingual after Injury

Having assessed the patient's language knowledge and use before injury, it will be important to examine the patient in the speech modes he or she was involved in prior to injury.

(i) The monolingual speech mode. In the sessions examining the monolingual speech mode, it will be important to deactivate the language not being tested. To do this, the patient will have to be tested in EACH of the two languages (if both were used monolingually) at DIFFERENT TIMES and with different examiners WHO DO NOT KNOW THE OTHER LANGUAGE AT ALL. Thus, in each case, the patient will understand that he or she is facing a monolingual interlocutor and can therefore only use one language. We should note that in order to simulate the monolingual mode, many examiners "pretend" not to know the other language. This is quite inappropriate as the pretense is rarely foolproof and never lasts very long; the consequence is that the data obtained are usually ambiguous as they emanate from a conversation where the speech mode has changed from being monolingual to bilingual.

Keeping in mind the knowledge, use, and functions of the languages
prior to injury, it will now be possible to assess the impact of the injury on each of the two languages when they are used monolingually. Of particular interest will be amount and type of language loss as well as the kind of interferences that now occur: are these different from those prior to injury? It will also be necessary to determine if the patient can keep his or her two languages separate in these monolingual testing situations; change of base language or actual code-switching with a monolingual examiner will be a sure sign that the mechanism that allows bilinguals to deactivate one language, when speaking the other, has been affected.

(ii) The bilingual speech mode. If the patient also operated in the bilingual speech mode before injury, he or she will need to be examined in that particular mode. To do this, a testing situation will need to be set up such that the patient feels comfortable code-switching and borrowing during the examination. One way of doing this is to adjoin to a THIRD, bilingual, examiner (the first two were monolingual in either language A or B), some members of the patient’s family, or close friends with whom he or she code-switched and borrowed before injury. In this bilingual mode, one should study the appropriateness of language choice and the ability to code-switch and borrow. Questions that need to be answered are:

—Does the patient speak the “wrong” language to a bilingual family member or close friend?
—Does he or she mix language to the same extent as before?
—Are these mixes of the same type (code-switches, borrowings)?
—Are the code-switches still grammatically constrained?
—Do they belong to the same class: intersentential, intrasentential, single items, tags, etc.?
—Can the patient translate from one language to the other in the same way as he or she did before injury? etc.

Examining bilinguals in their various speech modes and determining the exact nature of the deficit in these modes should help us better understand bilingual aphasia, and more generally, the neurolinguistics of bilingualism.

CONCLUSION

A number of positive consequences will emerge from viewing the bilingual as a unique and specific speaker-hearer. First, it will encourage us to study the bilingual as a whole. We will no longer examine one of the bilingual’s languages without examining the other; rather we will study how the bilingual structures and uses the two languages, separately or together, to meet his or her everyday communicative needs. Second,
it will force us to use tests that are appropriate to the domains of language use: domains that involve mixed language will be tested in mixed language; domains requiring a monolingual speech mode will be tested monolingually. Care will also be taken not to give bilinguals batteries of tests that have little to do with their knowledge and use of the two languages. Third, this view will stimulate us to identify and control the speech mode bilinguals are in before recording or testing them. Fourth, it will force us to differentiate between the person who is in the process of becoming bilingual, and the one who has reached a stable level of bilingualism, whatever the ultimate level of proficiency attained in each of the two languages. Finally, this view will encourage us to study the bilingual as such and not always in comparison to the monolingual. We should always keep in mind that half the world’s population is bilingual and that using the monolingual as a yardstick is questionable.

REFERENCES