〔Article〕

Native Speaker Criteria in CPH Research:
Why Does It Matter?

Eiji Hashimoto

Abstract

This paper overviews the native speaker criterion used as a yard stick in language acquisition especially in the context of Critical Period Hypothesis（CPH）research. First, it identifies the contradiction that the criterion has been used extensively even though the validity of the criterion has been acknowledged to be infirm. Second, it refers the logic that is used to explain the difference between the L2 ultimate attainment and L1: the L2 speaker has multicompetence and the L1 speaker exhibits monocompetence. Third, three practical problems related to the use of the criterion are focused upon: definition of the native speaker, validity in testing, and switching from L2 to L1. Finally, an alternative criterion is discussed.

臨界期仮説の研究において，ネイティブスピーカーが第2言語話者の言語能力の判定基準として使われてきた。研究者の多くは，その基準を使用することに問題があると指摘しているにもかかわらず，未だにその傾向は大きく変わることがない。本稿では，その矛盾に関して次の4点を中心に考察する。1）臨界期仮説の研究におけるネイティブスピーカーの基準使用についての背景。2）ネイティブスピーカーと第二言語話者の言語能力の違い。3）基準使用に関する問題点。4）新たな基準の可能性。

Introduction

In the context of language acquisition, the history of CPH started with Penfield（1959）. He stated that a biological clock is attached in the child's brain and that allows him/her to have a specialized capacity to learn language through natural exposure. The age of nine was set as a time limit for cerebral flexibility that enables children to acquire language through direct input. After nine years of age, language is learned analytically and then achieving a native like proficiency would be more difficult. Lenneberg（1967）developed the idea, based on a biological and neurological explanation, by hypothesizing the lateralization of a hemisphere is a trigger to the critical period for language acquisition and suggested that it begins about the age of 2 and ends at puberty.

Since the late 1960's, after Penfield（1959）and Lenneberg（1967）had presented the notion of Critical Period（CP）to the field of language acquisition, this issue has been extensively debated and it is still not proven whether CP exists or not. Some researchers are opposed to the existence of CP（Bialystok & Hakuta, 1999; Birdsong, 1992, 2006, 2007; Birdsong & Molis, 2001; Bongaerts, 1999, 2003; Bongaerts, Mennen, & Van Der Slik, 2000; Bongaerts, van Summeren, Planken, & Schils, 1997; Colantoni, & Steele, 2006; Felix, 1985; Marinova-Todd, 2003; Marinova-Todd, Marshall, & Snow, 2000; Montrul & Slabakova, 2003; Van Boxtel, Bongaerts, & Coppen, 2005; White & Genesee, 1996）, others

11

Native Speaker Criteria in CPH Research : Why Does It Matter?

阪南論集　人文・自然科学編　　　　　　　　　　　　　　　Vol. 48 No. 1

are in favour of it（Coppieters, 1987; DeKeyser, 2000; Hyltenstam, 1988; Hyltenstam, & Abrahamsson, 2000, 2001, 2003a, 2003b; Johnson, & Newport, 1989; Long, 1990, 2005, 2007; Patkowski, 1980; Scovel, 1988, 2000, 2006）, and there are also researchers who admit that there are extremely talented language learners who pass the native speaker level in some linguistic areas although the number is very limited（Ioup, 1995; Ioup, Boustagui, Tigi, & Moselle, 1994; Moyer, 1999）.

As many researchers prefer the term sensitive period instead of CP, this phenomenon could neither have an all or nothing effect, nor cause abrupt onset-offset because language acquisition could also occur outside this period（Hyltenstam & Abrahamsson, 2001）. This characteristic is probably one of the factors that could make this research more complicated. Hyltenstam and Abrahamsson（2003a）summarize the past empirical findings as follows:

> While few researchers today would deny long-term advantages for child starters, especially after reviews have found no counter evidence to this, on child starters（Krashen, S., Long, M. H., and Scarcella, R.（1979）; Long, M. H.（1990）; Singleton, D.（1989; 2001）views differ as to whether these observations should be explained by biological scheduling, that is, by constraints imposed on the learner along with maturation, or by social/psychological factors.（pp. 539-540）

There is little consensus on CPH studies besides the above example inspite of decades of research quests and the controversy seems to never end. Hyltenstam and Abrahamsson（2003a）consider that this endless discussion is caused by the notion of *the native speaker*. They assert that empirical data in the field lacks validity and could jeopardize the foundation of the CPH research because there is no explicit criterion as to what is meant by native speaker proficiency.

Birdsong（2006, 2009）points out that in the literature the native speaker has been referred as the monolingual in the CPH research context and has been used interchangeably as a benchmark, and sometimes incorrectly, with the term L2 ultimate attainment. Moreover, the criterion seems to vary from the sample of the native speakers' performance to their intuitive competence which is then compared qualitatively or quantitatively with those of L2 speakers'.

There is a major contradiction here; researchers have been using the native speaker as a criterion in CPH research and at the same time criticizing its validity. Many researchers claim that this criterion should be reconsidered（cf. Bley-Vroman, 1989; Cook, V., 1992, 1999; Davies, A., 2003, 2004; Kecskes & Papp, 2000; Long, 1990; Munoz, 2011; Pavlenko, 1999; Piller, 2002; Schachter, 1990）. This paper reexamines the concept the native speaker as a yard stick in the contradiction. First, it focuses on differences between L1 and L2 ultimate attainment. Second, it refers to the factors that could be involved in the cause of the differences. Third, it reflects problems when the native speaker criteria are actually used in a research design. In conclusion, a discussion and suggestion for future research are made from the perspectives of validity and practicality.

Is L2 ultimate attainment the same as L1?

The literature confirms that the native speaker norm should be reconsidered（Bley-Vroman, 1989; Cook, 1992, 1999; Davies, 2003, 2004; Kecskes & Papp, 2000; Long, 1990; Pavlenko, 1999; Piller, 2002;

Munoz, 2011; Schachter, 1990). What implies is that there is no explicit criterion for native speaker proficiency, and there will be no point in discussing whether the L2 learner could reach the native level of language proficiency or not. To proceed with a discussion without having a logical reasoning problem, this paper considers that the native speaker criterion has been valid and properly used in research within a margin of error.

The most popular logic that has been used to criticize the deficiency of the native speaker norm is based on the assumption that L2 ultimate attainment is not the same as L1. Some studies acknowledge that young L2 learners could become quite fluent speakers with possibly a non-native accent (e.g., Flege, 1999, 2002; Flege, Frieda, & Nozawa, 1997; Guion, Flege, & Loftin, 2000; Piske, Mackay, & Flege, 2001) and others maintain that their L2 ultimate attainment is different from the L1 (Birdsong, 2009; DeKeyser, 2000; Flege, 1999; Hyltenstam, 1992; Hyltenstam & Abrahamsson, 2000, 2003b; Sorace, 1993, 2003). Some studies assert that the proficiency of very young starters and high proficient bilinguals would be no different from that of the native speakers (Flege, 1995; Mack, 1989; Montrul, 2002, 2011a; Pallier, Bosch, & Sebastia´n-Galle´s, 1997; Wong Fillmore, 1991).

For example, Flege (1995) has detected the presence of a foreign accent in the sound of /i/, / ε /, and / α / between early bilinguals of Italian and L1 monolingual English speakers, and he concluded that the difference results from the degree of perceived similarities between the two languages. Furthermore, Mack (1989) investigated French bilinguals and L1 English speakers' ability to identify and discriminate the synthetic English sounds of /i/ and /ɪ/. Results were that they could discriminate the difference between the vowels but there was difference in the location of phoneme boundaries and he reached the conclusion that their perception of the vowels is very similar with that of the L1 speakers' but not the same. In a similar research framework to Mack's study, Pallier, Bosch, and Sebastia´n-Galle´s (1997) studied bilingual speakers of Catalan and Spanish compared with L1 Spanish speakers. They assessed the identification and discrimination of the synthetic Spanish sound of /e/ and / ε / by those speakers and concluded that the influence of L1 is so persistent, even if exposure to L2 language was begun at an early age, that the malleability to a new phonetic system is considered to be severely limited.

In bilingual studies there are also similar results reported. Wong Fillmore (1991) conducted a nationwide study on bilingual children in the U.S.A. to find out how timing and social factors affect the retention and continued use of their primary languages and the development of their second language when they settled in a new English speaking environment. He concluded as follows:

> Social and psychological factors have been implicated in the case of adult language learners, whereas situational factors have been critical in the case of children. Fossilized interlanguages are very likely to develop in the language-learning situations that we find in many schools with large enrollments of immigrant and refugee students. In the classrooms of such schools, non-English speakers frequently outnumber English speakers. In fact, except for their teachers, the learners may have little contact with people who know the language well enough to help them learn it. In any event, the language learners spend a lot more time talking with one another than they do with their teachers, and the English they hear most often is the imperfect varieties spoken by classmates rather than the more standard varieties spoken by their teachers. That

13

Native Speaker Criteria in CPH Research : Why Does It Matter?

阪南論集　人文・自然科学編　　　　　　　　　　　　　Vol. 48 No. 1

being the case, the input they base their language learning on being the speech of learners like themselves, is not altogether representative of the target language. Not surprisingly, language learning based on such input is neither perfect nor complete.（p. 345）

Results of the research support the claim made by Hyltenstam（1992, 1988）suggesting that "an early Age of Onset（AO）may be a necessary although not sufficient requirement for nativelike ultimate attainment"（p. 364）.

Furthermore, Montrul（2002）examined the effect on ultimate attainment in Spanish by age of onset of English and Spanish bilingual speakers. The existence of incomplete acquisition and attrition of tense/aspect morphology in Spanish were recognized. Tests consisted of an oral production, a written completion and two meaning-interpretation tasks. Results showed that there was a significant statistical difference between both of the simultaneous bilingual and the early child L2 groups and L1 Spanish speakers, and explained that "morpho-phonological spell-outs and semantic features of functional categories are affected by incomplete acquisition and language loss"（p. 39）. She concluded that "the earlier the age of onset of bilingualism and the more intense the exposure to the sociolinguistically dominant language, the more incomplete the adult grammar may turn out to be, suggesting that an early age of onset of a critical period might be a necessary but not a sufficient condition for convergent language acquisition"（p. 61）.

In other perspectives, Ekberg（1998）investigated frequency differences in the use of specific structures and vocabulary between the L1 Swedish speaker and the L2 early learner. Results showed that significant difference was identified in sentence connectors, presentation, pseudo-coordination expressing progressive aspects, and complex predicates. In a similar study, McDonald （2000）investigated the grammatical proficiency of native English speakers and early L2 speakers of Vietnamese and Spanish. Results showed that there was no statistical difference in a grammatical judgment test between the Spanish native speakers and the English native speakers in contrast to the Vietnamese speakers whose English was judged not to be identical with the native English speakers'. The Vietnamese speakers had problems with such structures as articles and plurals that are markedly different from English. It was concluded that the result could more likely be attributed to their L1 than the amount and quality of input because those early Vietnamese speakers had been exposed to a massive English input at schools where English was spoken since the age of 5（cf. Hyltenstam & Abrahamsson, 2001; Butler, 2000）.

There seem to be cases where even young starters have some difficulty passing the native speaker level, not only in one specific area but also in several areas. Most of the researchers consider L1 attrition and the amount and quality of input play a key role for ultimate acquisition and L2 ultimate attainment is not an automatic phenomenal outcome（Harley & Wang, 1997）. Other than a biological factor, there could also be influential social and psychological factors involved in the process of second language acquisition. In the next section, the logic of why L2 ultimate attainment is different from L1 is discussed.

What is the cause of the difference?

As described in the previous section, the majority of researchers agree that L2 ultimate attainment fundamentally differs from L1. The logic behind the formulation is that the difference is caused by the multicompetence of the L2 speaker. The concept of multicompetence was introduced by Cook (1991) and it originates from Univversal Grammar (UG) theory (Cook, 1992; Cook & Newson, 2007).

Cook, V. (1991) introduced the concept of mulicompetence "the compound state of a mind with grammars" (p. 112) and distinguished the monocompetence of L1 "the state of the mind with only one grammar (ibid). Researchers such as Piller (2002), Kecskes and Papp (2000), and Pavlenko (1999) have the same standpoint that they are fundamentally different so that two different kinds of competences would not produce the same performance. According to Cook (1999), the proficiency of the L2 speaker is viewed as different rather than deficient. He explains that "multicompetence covers the total language knowledge of a person who knows more than one language, including both L1 competence and the L2 interlanguage (p. 190), and also suggests that multiconpetence does not mean two monocompetences exist in one mind (Cook, 1992). Grosjean (1989) describe a prevalent view on bilingualism as follows:

> 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. (p. 4)

Grosjean (1989) claims that this belief unfortunately spread among many researchers and educators in the language science and that the monolingual speaker has always been used as a benchmark.

The notion of mono vs. multicompetence branched from UG theory (Cook & Newson, 2007). UG theory emphasizes that the accessibility to a parameter is considered to be a key to acquire language: natural acquisition is possible as long as the parameter is active. These remarks indicate that there is a critical time limit for both the L1 and L2 speaker to access the parameter. As far as L2 acquisition is concerned, Age of Acquisition (AOA) is considered as an important factor to make a fundamental difference between the early and late L2 learner's performances.

Unfortunately, there are very few direct formulations to explain the difference between L2 ultimate attainment and L1 except the notion of mono vs. multicomptence. Although UG theory is deeply concerned with multicompetence, neither accounting for the development of UG in relation to multicompetence is the keystone of this paper nor is there sufficient room in this article so that we will move to focus on problems when the native speaker criteria are used in research.

What kind of problems does the native speaker criteria have?

The concept of the native speaker has been indispensable in the field of linguistics and its

Native Speaker Criteria in CPH Research : Why Does It Matter?

阪南論集　人文・自然科学編　　　　　　　　　　　Vol. 48 No. 1

fundamental importance is unlikely to change. The difficulty in the use of the concept remains in its illusory definition. The definition of the native speaker and its implications vary according to different disciplines (Coulmas, 1981). Raffler-Engel (1981) describes how this ideal mythical creature has been used as follows:

> Anthropological linguists were aware of the fact that speech communities are not homogeneous but did not worry about increasing their subject population for statistical purposes. They simply layered their data according to one average old informant, one average middle aged informant, and one average young informant.
>
> 　The sociolinguists divided native speakers by social class and even realized that each social class had more than one speaker. Eventually the native speaker was also endowed with a cultural background which manifested itself in the ethnography of speaking.
>
> 　The average native speaker was always male. The average woman entered the subject pool only recently. She is generally still so average that be she young and middle class or old and lower class she can be compared to any type of average man. She is also so average that her verbal behavior can be examined without regard to the social situation in which it takes place. (p. 300)

　Kravchenko (2010), Coulmas (1981) and Singh (1998, 2006) argue that a native speaker criterion has a number of methodological problems and has been used as a norm for a practical reason even though researchers acknowledge the problems. Some of the researchers such as Long (1993) contend that the native speaker norm should be reconsidered because it could increase a margin of error, influence validity and reliability and as a result it could provide a different interpretation.

　In the field of CP research, there are three problems when the native speaker criteria are used. First of all, as with the other disciplines of linguistics, the definition of the native speaker and her/his language are a significant problem (Bachman, 1990; Davies, 1988; Lantolf & Frawley, 1985; Paikeday, 1985). Love (2010) describes the problem as follows:

> Languages are not natural objects, and they do not become natural objects by theoretical fiat. There has never been agreement as to what a language is or even how we might set about deciding what a language is. What counts as a language is a matter of cultural and political context and purposes, and these can vary without limit. As every undergraduate introduction to linguistics points out, there is no justification for distinguishing, say, Swedish and Danish as separate languages while treating Mandarin and Cantonese as variants of the same language. There is no consensus as to where one language ends and another begins, either in time or in space or in terms of social differentiation. (p.592)

　Love (2010) also refers to a real situation representing the difficulty of using the term native language in the Indian government census by quoting Pattanayak (1981). In the census the native language was defined as 1) "the language spoken by the individual from the cradle", 2) "parent tongue", and 3) "language ordinarily spoken in the household" (pp. 47-48). Even if the targeted

native speaker would be secured for experiments, careful attention still should be paid to the usage. Mufwene (1998) explains the difficulty of collecting reliable language data from the native speaker subject as follows:

> Consequently, for the purposes of collecting reliable date on particular languages, and insofar as indigenized varieties are recognized as valid for linguistic investigation, having informants who are proficient in a variety seems far more relevant than having native speakers. In some communities, being an adult native speaker of one language variety also amounts to being proficient in it; however, in some others this equation does not hold. (p. 121)

As Mufwene (1988) observes, a community could be more complex than it looks. It has to preserve geographical, economic, political, social, educational, cultural, ethnographical, and historical variables inside and those variables are intricately woven. It is more difficult now to find a community in which a target native language is used and is still eligible to represent a reliable source of information as a language norm because the world is more mobile than ever.

Secondly, switching from L2 to L1 is a problem to identify the native speaker. Technically speaking, it would be very difficult to distinguish young L2 learners who claim their L2 is a dominant language from those who claim their L2 is still the second language although their proficiency is exactly as the same as the native speaker's. Should the former be called the native speaker or the bilingual?

There are many claims in the literature that L2 could be replaced by L1 if conditions are met and the processed L2 should be regarded not as L2 but a continuation of L1 acquisition (Bialystok 1997; Bialystok & Hakuta, 1994; Harley, & Wang, 1997). Kouritzin (1999) reports the case of Lara who was an immigrant from Finland and had moved to Canada when she was two years old. She lived with her family in a Finnish community until at the age of six when the family moved to a large city to more integrate with the main stream Canadian culture where English is used most of the time. By the age of 18, she considered that her Finnish was replaced by English and English became L1. Jia and Aaronson (2003) urge that switching from L2 to L1 among young immigrants more likely occurs if AOA is before ten and the first language tends to be maintained after that age. They concluded that this was caused by lack of or no support from the community the young immigrants belong to, or their parents' educational policy was to encourage them to integrate with the main stream culture in a host country. Flege (1999) also asserts that a frequent tradeoff among immigrant children is recognized especially in the phonetic-phonological domain. According to his claim, young immigrants could master L2 without a foreign accent at the expense of L1 loss, as L1 and L2 phonetic-phonological systems unavoidably interact with each other in the process of acquisition. The evidence of the incomplete acquisition of L1 and mutual interaction are observed in many cases of heritage speakers whose proficiency varies from the state of complete lost to the native level (cf. Montrul, 2011b, 2010a; Montrul & Ionin, 2010b).

Finally, in the area of testing, quite a few researchers are against the use of the native speaker criteria (Alderson, 1980; Carroll, 1971, 1983; Davies, 1988, 2003; Jafarpur, 1996; Johnson, 1981; Long, 1993; Marinova Todd, Marshall, & Snow, 2000). For example, Long (1993) states that the research

instruments used in most of the CPH studies are not precise enough to measure language proficiency accurately. Davies（1988）refers five unreliable variables in measuring language proficiency: "the native speaker, cut-off, the criterion score, the test and the language"（p. 32）. The native speaker criterion is problematic; however, there are other factors that are associated with the problem.

First of all, Mrinova-Todd（2000）points out that the native speaker has accents and they could necessarily be perceived as the standard and criticizes that most of the studies designed to investigate L2 pronunciation failed to specify "either the exact margins of what is considered a standard accent in the target language or the degree of variability among native speakers"（p. 19）. Furthermore, she stated that the background information, perception of L2 subjects, and a judge's belief and experience in language acquisition could influence their decisions. Consequently, a native speaker could be judged as a non-native speaker in some area in the host country and vice-versa（cf. Long, 1993）.

Second, some researchers argue that there is a variation in the native speaker's proficiency （Bachman, 1990; Davies, 1988; Oiler & Spolsky, 1979; Paikeday, 1985; Stump, 1978）. Davies（2003） claims that the native speaker's proficiency is also developmental suggesting that "however close the match between our modeling of proficiency and the native speaker may be, we can never be sure that there is a complete match"（p. 173）. Furthermore, interesting data on native speaker proficiency provided by Bachman（1990）suggests that "there is growing evidence that native speakers perform neither uniformly well on tests of all aspects of language ability, nor uniformly better than do non-natives"（pp. 248-249）.

Third, a scoring method could play an important role on the effect of test performance（Bachman, 1990）. For instance, Crewe（1977）examined Singapore English quoted by Goh Poh Seng's novel *If We Dream Too Long*（1971）and commented that it could be very difficult to classify errors that were made by the non native speaker because the native speaker makes the same type of errors as well. This is a rather complicated case, however, unless a specific scoring guideline is provided and executed correctly as it is designed, validity would be absent in a test.

Lastly, selecting an irrelevant test has many implications to the native speaker criterion. It would be impossible to describe all the details of the native speaker's language proficiency and to measure a particular one designed by a researcher because there are many complicatedly compounded factors that affect performance in tests（Jafarpur, 1996）. Although it is acknowledged that the native speaker criterion has good face validity（Lado, 1978）and usefulness（Bachman, 1990）, the validity of a test should be carefully examined as to whether the target language proficiency that the test is designed to measure corresponds with the objective of a tester. Jafarpur（1996）points out that "scores on tests are not so much affected by the specific trait the tests intend to tap as influenced by the strategic competence of the testees"（p. 83）.

For example, Alderson（1980）examined the native and non-native speaker's performance on the cloze tests in which different variables are set up according to the purpose of each experiment. Results showed that the native speaker's performance was slightly better than the non-native speaker's. However, either the former failed to reach a perfect score or there was no clear distinction between them. It was concluded that:

attempts to use native speakers as criteria for non-native speakers, as in clozentropy or much criterion-referenced testing, are misguided. If native speakers vary in their ability to do cloze tests, then the value of using them as criteria is doubtful". (p. 75)

He considers that the cause of variance results from intelligence: intelligence is closely related to language proficiency (cf. Oller, 1978) and that normal distribution of intelligence naturally occurs in the native speaker's population and thus the native speaker's language proficiency also varies.

Four basic problems have been focused on when the native speaker criterion was used in a practical situation. More problems would be probably identified if details are cited. Now, the time is for a final question: Should the native speaker criterion be abandoned or not? In the next section, the question is discussed as the conclusion of this paper.

Discussion

The contradiction of the native speaker criterion is two holds: most CPH researchers use the criterion but at the same time they claim that it should be reconsidered because it does not have firm validity (Alderson, 1980; Carroll, 1971, 1983; Davies, 1988, 2003; Jafarpur, 1996; Johnson, 1981; Long, 1993; Marinova Todd, Marshall, & Snow, 2000). The reason why the standard has been used in studies is its practicality (Bachman, 1990; Lado, 1987).

Researchers such as Birdsong (2009), Cook (1992), and Munoz (2011) suggest the bilingual speaker or the highly proficient early L2 speaker as an alternative choice. The basic logic of the proposal is that the L2 speaker has multicompetence and the L1 speaker has monocompetence and therefore, the standard should be the multicompetent L2 proficiency. However, this new criterion has at least two issues to be considered before it is put in practice.

One is the definition of the bilingual speaker. Grosjean (1989) describes the bilingual speaker as follows:

> 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. (p. 4)

The problem of the bilingual speaker criterion is more complicated than the criterion of native speaker because proficiency varies much more in the case of the bilingual speaker. It is also important to know whether the bilingual speaker is "the person who is in the process of becoming bilingual", or "the one who has reached a stable level of bilingualism (Grosjean, 1989, p. 14). Defining

Native Speaker Criteria in CPH Research : Why Does It Matter?

阪南論集　人文・自然科学編　　　　　　　　　　　　　　Vol. 48 No. 1

the stage at which the person is located and measuring their proficiency would be a very difficult task. Grosjean（1989）refers to tests that have been used for the bilingual speaker as follows:

> 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 [sic] for the two languages or the DIFFERENT SOCIAL FUNCTIONS [sic] 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.（p. 4）

It would be challenging to write a test for the bilingual speaker because those needs and functions vary in individuals and they have more complicated educational, linguistically based, social backgrounds than the native speakers have, even in one community.

The other problem of the bilingual speaker criterion is the notion of switching: L2 becomes L1. This phenomenon has been empirically proved by many researchers（Bialystok, 1997; Bialystok & Hakuta, 1994; Flege, 1999; Harley & Wang, 1997; Kouritzin, 1999; Montrul, 2011b, 2010a; Montrul & Ionin, 2010b）. Munoz（2011）contends as follows:

> There is a sense, therefore, in which making a comparison between the proficiency of such language users in their dominant language and that of native speakers is not a matter comparing L2 attainment with an L1 baseline, because for such language users their dominant language for all practical purposes is ITSELF [sic] their L1.（p. 8）

Assuming this claim that the bilingual whose dominant language is the same in quality and proficiency as the native speaker's, would it be possible to say that using the bilingual speaker criterion would be the same as using the native speaker criterion. In other words, there are two types of L1 speakers: one is the monolingual native speaker and the other one is the bilingual speaker and their performance is identical.

In conclusion this paper suggests that it is not important whose L1 should be used as a criterion because once the monolingual native speaker learns a foreign language at school, technically they are no longer the monolingual native speaker but the multicompetent bilingual speaker. The world is already full of bilingual speakers（Grosjean, 1989）. The line between the native speaker and the fluent bilingual speaker is very fuzzy. Davies（2003）describes that "the distinction native speaker– non-native speaker . . . is at bottom one of confidence and identity"（p. 213）. What is more important is controlling variables between an experimental and control group. Long（1993）suggests, if the native speaker criterion is the only choice, that "Ns controls should be comparable in age, sex, education, and social class, and in any other way known to be sociolinguistically relevant in the language or dialect concerned"（p.125）. Where the native speaker or the bilingual speaker is used as a criterion, describing participants in detail would be one key and controlling variables is the other.

References

Alderson, J. C. (1980). Native and nonnative speaker performance on cloze test. *Language Learning*, 33(1), 59-76.

Bachman, L. F. (1990). Fundamental Considerations in Language Testing. Oxford, Oxford University Press.

Bialystock, E. (1997). The structure of age: In search of barriers to second language acquisition. *Second Language Research*, 13(2), 116-137.

Bialystok, E., & Hakuta, K. (1999). Confounded age: Linguistic and cognitive factors in age differences for second language acquisition. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 161–181). Mahwah, NJ: Erlbaum.

Bialystok, E., & Hakuta, K. (1994). *In other words: The Science and Psychology of Second-Language Acquisition*. New York: Basic Books.

Birdsong, D. (2009). Age and the end state of second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 401-424). Bingley: Emerald.

Birdsong, D. (2007). Nativelike pronunciation among late learners of French as a second language. In O.-S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emil Flege* (pp. 99–116). Amsterdam: Benjamins.

Birdsong, D. (2006). Age and second language acquisition and processing: A selective overview. *Language learning*, 56(S1), 9-49.

Birdsong, D., & Molis, M. (2001). On the evidence for maturational constraints in second-language acquisition. *Journal of Memory and Language*, 44, 235–249.

Birdsong, D. (1992). Ultimate attainment in second language acquisition. *Language,* 68(4), 706–755.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart & Winston.

Bley-Vroman, R. (1989). What is the logical problem of foreign language learning? In S. M. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 41-68). Cambridge, Cambridge University Press.

Bongaerts, T. (2003). Effets de l'ˆage sur l'acquisition de la prononciation d'une seconde langue. *Acquisition et Interaction en Langue E´trang`ere*, 18, 79–98.

Bongaerts, T., Mennen, S., & Van Der Slik, F. (2000). Authenticity of pronunciation in naturalistic second language acquisition: The case of very advanced late learners of Dutch as a second language. *Studia Linguistica*, 54(2), 298–308.

Bongaerts, T. (1999). Ultimate attainment in L2 pronunciation: The case of very advanced late L2 Learners. In D. Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp. 133–159). Mahwah, NJ: Erlbaum.

Bongaerts, T., van Summeren, C., Planken, B., & Schils, E. (1997). Age and ultimate attainment in the pronunciation of a foreign language. *Studies in Second Language Acquisition*, 19(4), 447–465.

Butler, Y. G. (2000). The age effect in second language acquisition: Is it too late to acquire nativelevel competence in a second language after the age of seven? In Y. S. Oshima-Takane & H. Sirai (Eds.), *Studies in language sciences* (pp. 159-169). Tokyo: The Japanese Society for Language Sciences.

Carroll, B. J. (1983). Issues in the testing of language for specific purposes. In A. Hughes & D. Porter (Eds.), *Current Developments in Language Testing* (pp. 109-114). London: Academic Press.

Carroll, J. B. (1971). Language and cognition: current perspectives from linguistics and psychology. In Laffey, J. F. Laffey & R. Shuy (Eds), *Language Differences. Do They Interfere?* (pp.173-185). Newark: International Reading Association.

Colantoni, L., & Steele, J. (2006). Native-like attainment in the L2 acquisition of Spanish stop-liquid clusters. In C. A. Klee & T. L. Face (Eds.), *Selected proceedings of the 7th conference on the acquisition of Spanish and Portuguese as first and second languages* (pp. 59–73). Somerville, MA: Cascadilla Press.

Cook, V., & Newson, M. (2007). *Chomsky's Universal Grammar: An introduction* (3rd ed.). Oxford: Blackwell.

Cook, V. (1999). Going beyond the native speaker in language teaching. *TESOL Quarterly*, 33(2), 185–209.

Cook, V. (1992). Evidence for multicompetence. *Language Learning* , 42, 557-591.

Cook, V. (1991). The poverty of the stimulus argument and multicompetence. *Second Language Research*, 7(2), 103-117.

Native Speaker Criteria in CPH Research : Why Does It Matter?

阪南論集　人文・自然科学編

Vol. 48 No. 1

Coppieters, R. (1987). Competence differences between native and near-native speakers. *Language*, 63(3), 544-573.

Coulmas, F. (1981). Introduction: The concept of native speaker. In F. Coulmas (Ed.), *A festschrift for native speaker* (pp.1-25). The Hague: Mouton Publishers.

Crewe, W. (1977). Singapore English as a non-native dialect. In W. Crewe (Ed.), The English Language in Singapore (pp. 96-119). Singapore: Eastern Unversities Press.

Davies, A. (2003). *The native speaker: Myth and reality*. Clevedon: Multilingual Matters Ltd.

Davies, A. (2004). The native speaker in applied linguistics. In A. Davies & C. Elder (eds.), *The handbook of applied linguistics* (pp. 431–450). Malden: Blackwell Publishing.

Davies, A. (1988). Operationalizing uncertainty in language testing: An argument in favor of content validity. *Language Testing,* 5, 32–48.

DeKeyser, R. M. (2000). The robustness of critical period effects in second language acquisition. *Studies in Second Language Acquisition*, 22, 499–533.

Ekberg, L. (1998). Regellampning kontra lexikonkunskap i svenskan hos invandrarbarn i Malmo. In P. Q. J. Moller, A. Holemen, & J. N. Jorgensen (Eds.), Nordiske sprog som andetsprog (pp.99–116). Institut of humanistiske fag: Danmarks Larehojskole.

Felix, S. (1985). More evidence on competing cognitive systems. *Second Language Research*, 1, 47–72.

Fishman, J. (1965). Who speaks what language to whom and when? *Linguistique*, 2, 67-88.

Flege, J. (2002). Interactions between the native and second-language phonetic systems. In P. Burmeister, T. Piske & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp.217-244). Trier: Wissenschaftlicher Verlag.

Flege, J. (1999). Age of learning and second language speech. In Birdsong (Ed.), *Second language acquisition and the critical period hypothesis* (pp.101–131). Mahwah, NJ: Lwqrence Erlbaum Associates.

Flege, J., MacKay, I., & Meador, D. (1999). Native Italian speakers' perception and production of English vowels. *Journal of the Acoustical Society of America*, 106(5), 2973–87.

Flege, J., Frieda, E., & Nozawa, T. (1997). Amount of native language (L1) use affects the pronunciation of an L2. *Journal of Phonetics*, (25)2, 169–186.

Flege, J. (1995). Second-language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: theoretical and methodological issues* (pp. 233–272). York: Timonium.

Grosjean, F. (1989). Neurolinguists, beware- the bilingual is not 2 monolinguals in one person. *Brain and Language*, 36(1),3-15.

Guion, S., Flege, J., & Loftin, J. (2000). The effect of L1 use on pronunciation in Quichua–Spanish bilinguals. *Journal of Phonetics*, 28(1), 27–42.

Harley, B., & Wang, W. (1997). The critical period hypothesis: Where are we now? In J. F. Kroll & A. M. B. de Groot (Eds.), *Tutorials in bilingualism: Psycholinguistic perspectives* (pp.19-51). London: Lawrence Erlbaum Associates.

Hyltenstam, K., & Abrahamsson, N. (2003a). Maturational constraints in SLA. In C. Doughty & M. H. Long (Eds.), The handbook of second language acquisition (pp. 539–588). Malden, MA: Blackwell Pub.

Hyltenstam, K., & Abrahamsson, N. (2003b). Age of onset and ultimate attainment in near-native speakers of Swedish. In K. Fraurud & K. Hyltenstam (Eds), *Multilingualism in global and local perspectives* (pp. 319–340). Stockholm Rinkeby Stockholm: Centre for Research on Bilingualism, Stockholm University, and Rinkeby Institute of Multilingual Research.

Hyltenstam, K., & Abrahamsson, N. (2001). Age and L2 learning: The hazards of matching practical "implications" with theoretical "facts." *TESOL Quarterly*, 35, 151–170.

Hyltenstam, K., & Abrahamsson, N. (2000). Who can become native-like in a second language? All, some, or none? On the maturational controversy in second language acquisition. *Studia Linguistica*, 54(2), 150–166.

Hyltenstam, K. (1992). Non-native features of near-native speakers: On the ultimate attainment of childhood L2 learners. In R. J. Harris (Ed.), *Cognitive processing in bilinguals* (pp.351-368). New York: Elsevier.

Hyltenstam, K. (1988). Lexical characteristics of near-native second-language learners of Swedish. *Journal of Multilingual and Multicultural Development*, 9, 67-84.

Ioup,G. (1995). Evaluating the need for input enhancement in post-critical period language acquisition. In Singleton & Z. Lengyel (Eds.), The age factor in second language acquisition (pp.95–123). Clevedon:

Multilingual Matters.

Ioup, G., Boustagui, E., Tigi, M., & Moselle, M. (1994). Reexamining the critical period hypothesis: A case study of successful adult SLA in a naturalistic environment. *Studies in Second Language Acquisition*, 16(1), 73–98.

Jafarpur, A. (1996). Native speaker performance validity: In vain or for gain? *System*, 24(1), 83–95.

Jia, G., & Aaronson, D. (2003). A longitudinal study of Chinese children and adolescents learning English in the United States. *Applied Psycholinguistics*, 24(1), 131-161.

Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology*, 21, 60-99.

Johnson, R. K. (1981). Questioning some assumptions about cloze testing. In J. Read (Ed.), *Directions in language testing* (pp. 117-206). Singapore: Singapore University Press.

Kecskes, I., & Papp, T. (2000). *Foreign language and mother tongue*. New Jersey: Lawrence Erlbaum Associates.

Kouritzin, S. G. (1999). *Face[t]s of first language loss*. Mahwah, NJ: Lawrence Erlbaum Associates.

Krashen, S., Long, M., & Scarcella, R. (1979). Age, rate and eventual attainment in second language acquisition." *TESOL Quarterly*, 13(4), 573-582.

Kravchenko, A. (2010). Native speakers, mother tongues and other objects of wonder. *Language Sciences*, 32(6), 677-685.

Lado, R. (1978). Scope and limitation of interview-based language testing: are we asking too much of the interview? In J. Clark (Ed.), *Direct testing of speaking proficiency: Theory and application* (pp. 115-128). Princeton: ETS.

Lantolf, J. P., & Frawley, W. (1985). Oral proficiency testing: a critical analysis. *The Modern Language Journal*, 69, 337-345.

Lenneberg, E. (1967). Giological foundations of language. New York: John Wiley.

Long, M. H. (2007). *Problems in SLA*. Mahwah, NJ: Lawrence Erlbaum Associates.

Long, M. H. (2005). Problems with supposed counter-evidence to the critical period hypothesis. *IRAL*, 43, 287–317.

Long, M. H. (1993). Second language acquisition as a function of age: Research findings and methodological issues. In K. Hyltenstam & A. Viberg (Eds.), *Progression and regression in language* (pp.196-221). Cambridge:Cambridge University Press.

Long, M. H. (1990). Maturational constraints on language development. *Studies in Second Language Acquisition*, 12(3), 251–285.

Love, N. (2010). The native speaker and the mother tongue. *Language Sciences*, 32(6), 589-593.

Mack, M. (1989). Consonant and vowel perception and production: Early English-French bilinguals and English monolinguals. *Perception & Psychophysics*, 46, 187–200.

MacKay, I., Meador, D., & Flege, J. (2001). The identification of English consonants by native speakers of Italian. *Phonetica*, 58(1–2), 103–125.

Marinova-Todd, S. H. (2003). *Comprehensive analysis of ultimate attainment in adult second language acquisition*. Cambridge, MA: Harvard University. (ProQuest information and learning company UMI Microform number 3100151).

Marinova Todd, S., Marshall, D., & Snow, C. (2000). Three misconceptions about age and L2 learning. *TESOL Quarterly*, 34(1), 9-34.

McDonald, J. (2000). Grammaticality judgments in a second language: Influences of age of acquisition and native language. *Applied Psycholinguistics*, 21(3), 395-423.

Montrul, S., & Slabakova, R. (2003). Competence similarities between native and near-native speakers: An investigation of the preterite-imperfect contrast in Spanish. *Studies in Second Language Acquisition*, 25, 351-398.

Montrul, S. (2011a). Multiple interfaces and incomplete acquisition. *Lingua*, 121(4), 591-604.

Montrul, S. (2011b). The linguistic competence of heritage speakers. *Studies in Second Language Acquisition*, 33(2), 155-161.

Montrul, S. (2010a). How similar are adult second language learners and Spanish heritage speakers? Spanish clitics and word order. *Applied Psycholinguistics*, 31(1), 167-207.

Montrul, S., & Ionin, T. (2010b). Transfer effects in the interpretation of definite articles by Spanish heritage speakers. *Bilingualism*, 13(4), 449-473.

Montrul, S. (2002). Incomplete Acquisition and Attrition of Spanish Tense/Aspect Distinctions in Adult Bilinguals.

Native Speaker Criteria in CPH Research : Why Does It Matter?

阪南論集　人文・自然科学編　　　　　　　　　　　　　　　Vol. 48 No. 1

Bilingualism, 5 (1), 39-68.

Montrul, S. (2001). First-language-constrained variability in the second-language acquisition of argument-structure-changing morphology with causative verbs. *Second Language Research*, 17 (2), 144-194.

Moyer, A. (1999). Ultimate attainment in L2 phonology. The critical factors of age, motivation, and instruction. *Studies in Second Language Acquisition*, 21 (1), 81–108.

Mufwene, S. (1998). Native speaker, proficienct speaker and norms. In R. Singh (Ed.), *The native speaker: Multiple perspectives* (pp.111-123).New Delhi: Sage Publications.

Munoz, C. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, 44 (1),1-35.

Oller, J., & Spolsky, B. (1979). The test of English as a foreign language. In B. Spolsky (Ed.), *Advances in language testing series 1: Some major tests* (pp. 92-100). Arlington: CAL.

Oller, J. (1978). How important is language proficiency to IQ and other educational tests? In J. Oller & K. Perkins (Eds.), *Language in education: testing the tests* (pp. 1-15). Rowley, Mass: Newbury House.

Paikeday, T. M. (1985). *The native speaker is dead?* Toronto: Paikeday Pub. Inc.

Pallier, C., Bosch, L., & Sebastia´n-Galle´s, N. (1997). A limit on behavioral plasticity in speech perception. *Cognition*, 64, B9–B17.

Papp, S. (2000). Stable and developmental optionality in native and non-native Hungarian grammars. Second Language Research 16 (2), 173-200.

Patkowski, M. (1980). The sensitive period for the acquisition of syntax in a second language. *Language Learning*, 30, 449–472.

Pattanayak, D.P. (1981). *Multilingualism and mother tongue education*. Delhi: Oxford University Press.

Pavlenko, A. (1999). New approaches to concepts in bilingual memory. *Bilingualism*, 2 (3), 209-230.

Penfield, W. (1959). *Speech and brain-mechanisms*. New York: Athenaeum.

Piller, I. (2002). Passing for a native speaker: Identity and success in second language learning. *Journal of Sociolinguistics*, 6 (2), 179–206.

Piske, T., MacKay, I., & Flege, J. (2001). Factors affecting degree of foreign accent in an L2: A review. *Journal of Phonetics* 29 (2), 191–215.

Raffler-Engel, W. (1981). The native speaker in his new fond body. In F. Coulmas (Ed.), *A festschrift for native speaker* (pp. 299-304). The Hague: Mouton Publishers.

Schachter, J. (1990). On the issue of completeness in second language acquisition. *Second Language Research*, 6 (2), 93-124.

Scovel, T. (2006). Age, acquisition, and accent. In C. Abello-Contesse, R. Chac´on-Beltr´an, M. D. L´opez-Jim´enez & M. M. Torreblanca-L´opez (Eds.), *Age in L2 acquisition and teaching* (pp. 31-48). Bern: Peter Lang.

Scovel, T. (2000). A critical review of the critical period hypothesis. *Annual Review of Applied Linguistics*, 20, 213–223.

Scovel, T. (1988). *A time to speak: A psycholinguistic inquiry into the critical period for human language*. Rowley, MA: Newbury House.

Singleton, D. (2001). Age and second language acquisition. *Annual Review of Applied Linguistics*, 21: 77-89.

Singleton, D. (1989). *Language Acquisition: The Age Factor*. Clevedon: Multilingual Matters.

Singh, R. (2006). Native speaker. In K. Brown & A. Anderson (Eds.), *Encyclopedia of language & linguistics* (2nd ed.) (pp. 489-492). Boston: Elsevier.

Singh, R. (1998). *The Native Speaker: Multiple Perspectives*. New Delhi: Sage Publications.

Sorace, A. (2003). Near-nativeness. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 130-151). Malden, MA: Blackwell.

Sorace, A. (1993). Incomplete vs. divergent representations of unaccusativity in non-native grammars of Italian. *Second Language Research*, 9 (1), 22-47.

Stump, T. A. (1978). Cloze and dictation tasks as predictors of intelligence and achievement scores. In J. W. Oiler & K. Perkins (Eds.), *Language in education. testing the* tests (pp. 36-64). Rowley: Newbury House.

Thiery, C. (1978). True bilingualism and second language learning. In D. Gerver & H. Sinaiko (Eds.), *Language interpretarion and communication*. New York: Plenum.

Van Boxtel, S., Bongaerts, T., & Coppen, P. A. (2005). Native-like attainment of dummy subjects in Dutch and the role of the L1. *IRAL*, 43, 355–380.

White, L., & Genesee, F. (1996). How native is near-native? The issue of ultimate attainment in adult second language acquisition. *Second Language Research*, 12(3), 233-265.

Wong Fillmore, L. (1991). When learning a second language means losing the first. *Early Childhood Research Quarterly*, 6, 323-346.