One Language or Two?
Real and Perceived Identification and Differentiation
of Language

Masatsugu MATSUO

Institute for Peace Science, Hiroshima University

SUMMARY

A separate language unique to an ethnonational group has significant role to play in an ethnonational movement and/or conflict. The language can be a symbol of the separate identity of the group. The language can be an important issue of the movement or conflict when it is excluded from the status of an official language. If there is a certain objective criterion by which the separateness of a language can be determined, then the task of those interested in the function of language in ethnonational conflicts will be to explore how the language is used for political purpose. There is, however, no such handy criterion. The identification and differentiation of languages are often made on non-linguistic, political grounds.

The present paper attempts, as a first step, to clarify the ways or patterns of identification and differentiation of languages. It provides a basis for better understanding of how a "separate" language is produced and is exploited for political purposes, especially for ethnonational purposes.

Introduction

It is often the case that, whether violent or not, ethnic (or ethnonational) conflict involves the identity of a language. If a separate language is viewed as evidence of the existence of a separate ethnic group or nation, then it is a serious question whether two related languages (or language varieties) are the same language or two different languages. It is, however, a very difficult question to answer. As Bernard Comrie argues:

> [It] is very difficult or impossible in many cases to decide whether two related speech varieties should be considered different languages or merely different dialects of the same language. […] but these decisions have often been made more on political and social grounds rather than strictly linguistic grounds (Comrie 1990: 2).

To understand the role of language in ethnic conflict, therefore, it is important to explore why there is such a difficulty. The present paper is an attempt to theoretically clarify the ways and patterns of language identification and differentiation. In order to achieve this, we should distinguish the reality and the perception of the relationship between two languages. It is because as Comrie suggests, the identification and differentiation depend largely upon political factors rather than purely linguistic factors. As Gianrenzo Clivio said: "[F]rom a strictly linguistic view, […] a language is a dialect that has an army and a navy and an air force; that is the only difference really from a linguistic point of view" (cited in Steinberg 1987: 199). The present paper tries to show the difficulty lies both in the very nature of human language and in the patterns of perception of language difference and identity.

1 Basic Assumption

Suppose first that there are only two languages to be considered. and let us refer to them as language A and language B, respectively. We will not examine cases where three or more languages are involved, because we think that it is best to begin with the simplest case.

Though we use the term language here, it should be understood as meaning "a possibly separate language," "what can be regarded as a separate language" or "a

dialect." A more precise term would be "language variety." But for simplicity's sake, we will use "language" instead.

Next we assume that the "languages" in question have something in common. In other words, we assume that each of the languages consists of two parts: that is, components unique to it and shared components (or parts). Following Noam Chomsky, we assume here that the relationship is always valid and universal, as to the relationship between any pair of languages, as far as they are human languages. The common elements may be "elements or properties of all human languages not merely by accident, but by necessity - of course […] biological […] necessity" (Chomsky 1975: 29). Or they may such elements as syntactic rules, lexical elements, and phonological elements. Then, the relationship in reality can be represented as Figure 1.

Figure 1 Basic Relationship
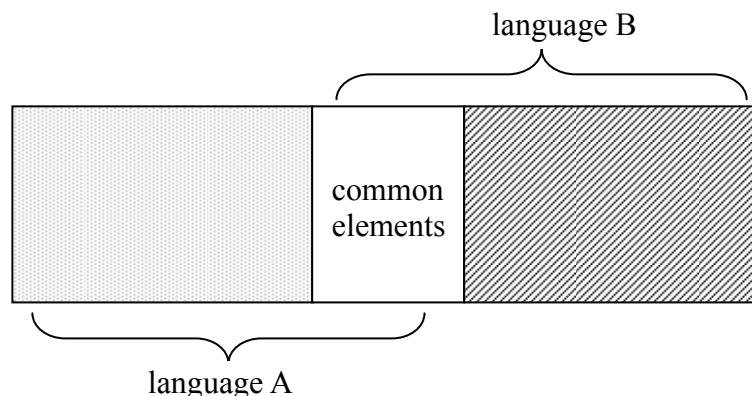
Given this basic realty, we can derive various perception patterns which focus only upon a certain aspect(s) of the relationship. The perceived relationship does not contradict the reality, that is, they are not necessarily false, but it does not represent the whole of the actual relationship. It is only a partial representation of reality in most of the cases.

Of such perceptions, we will first take up those ways of perception which differentiate the two languages. Next we proceed to cases in which the two are identified, or more precisely, cases where the two are regarded as belonging to one common language. Needless to say, there are cases in which we have difficulty in making such a distinction.

2 Differentiation

When the mutual intelligibility of the two languages, A and B, is rather low, it is natural that they should be considered to be two separate languages. In fact, there are actually two cases for this kind of differentiation. First, in some case, the existence of the common elements in our assumption is not perceived at all. Accordingly, the two languages are viewed as having nothing in common as in Figure 2. (Even in these cases, we should allow of some exceptions such as a small number of loan words). Of course, the distinction here is a matter of perception.

Figure 2 Differentiation with Nothing in Common
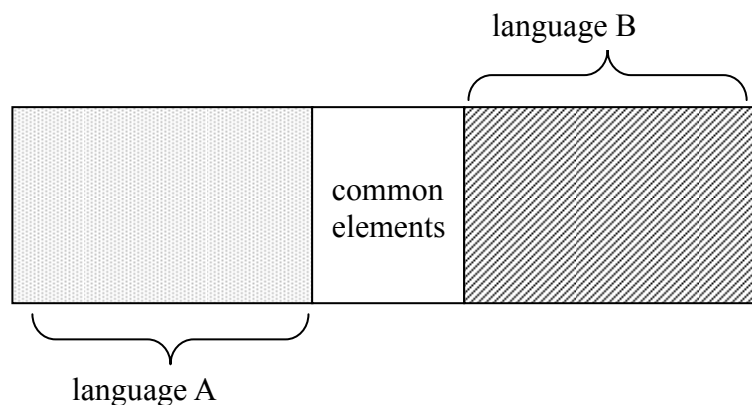
Secondly, in other cases, the existence of the common elements is clearly perceived. In this perception, the two languages A and B are distinct and separate with some thing common as is illustrated in Figure 3.

Though this perception is much closer to reality, the distinction is a matter of perception in this case again. It is usually not clear how much of the common part is included into each of the languages in question. For example, the language A may be perceived to include only a small part of the common elements. In this way, the perception of the languages is not always symmetric.

**Figure 3 Differentiation with Something in Common**
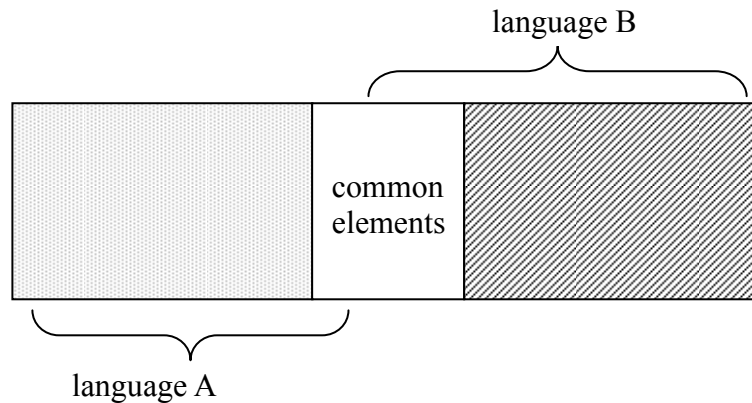
At first sight, such distinction seems quite a simple matter. But, the question remains open what degree or extent of commonness (to borrow a linguistic term, mutual unintelligibility or mutual intelligibility) is sufficient for such differentiation. And in fact, such a question-begging situation is not rare. For example, as to the Nordic languages, it is said that "Norwegian is Danish spoken in Swedish" (Haugen 1990: 151). With increasing mutual intelligibility or increasing amount of common elements, it may be assumed that it becomes more and more difficult to separate the two languages.

In the real world, however, this assumption does not necessarily hold. As is shown in Figure 4, very great commonness (or mutual intelligibility) does not necessarily prevent people from perceiving or claiming that the two are distinct languages. It is due to the great variation of the perceptions of the language A , for example, shown in Figure 4 (2), although the reality is like that given in (1).

Arguments for the existence of Scots are a case in point. Some argue that so-called Scottish English (which is generally regarded as a geographical dialect of English) is a language which is distinct from English and should be properly called "Scots" (Aitkin 1990: 76, McCllure et al 1980). If Scotland were to win independence from the United Kingdom as some dreamed of, and are still dreaming of, Scots might become one of the co-official languages of the independent Scotland together with Scottish Gaelic and (Standard) English.

Figure 4 Differentiation with Much in Common

(1) reality

language B

common
elements

language A

(2) perception

language B

common
elements
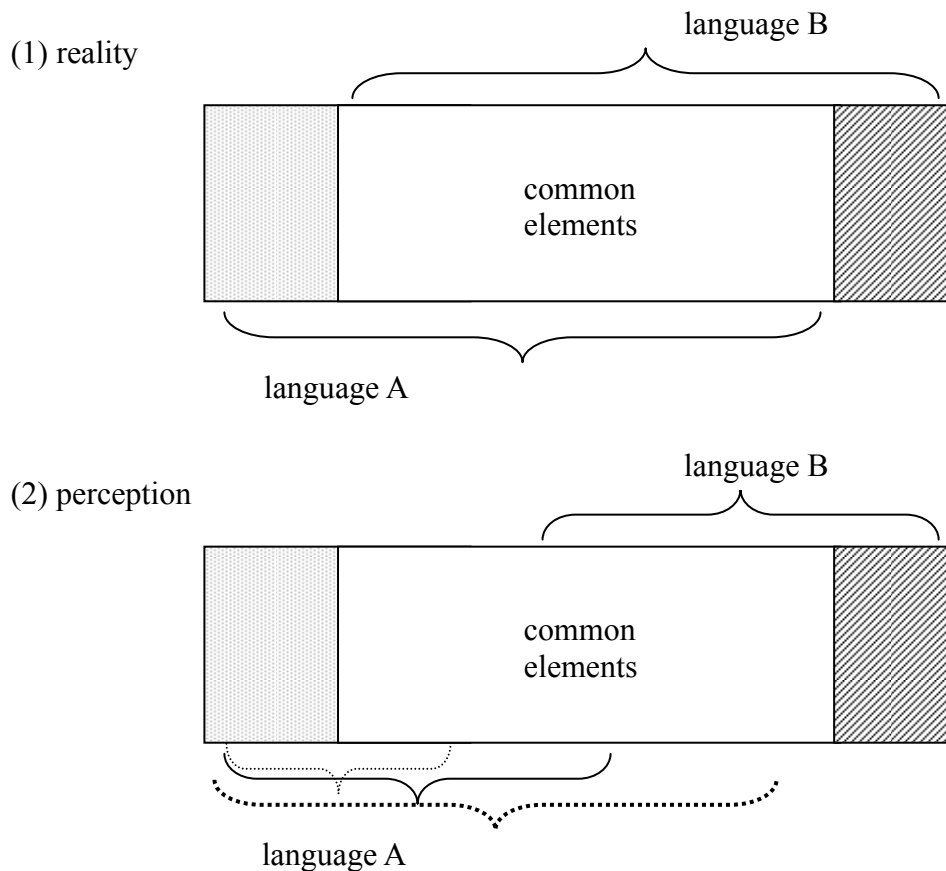
language A

There is another type of differentiation, which we can call asymmetric or partial differentiation. In this case, as is shown in Figure 5, one language, say, A, is perceived to be a language which is made up of unique parts and common parts according to our assumption.

On the contrary, the other language, B in this case, is not regarded as language. It is regarded as non-language at worst, and as a defective language at best. History provides us with many such examples.

For example, the Greek word "Bárbaros" originally meant stammering unintelligible speech to the Greeks. "Barbarous" languages of non-Greek people were not languages at all to the Greeks. In another example, a Dutch official is said to have compared Papiamentu / Papiamento, a creole language spoken now by some 300,000 in the island of Curaçao and others, to "the cackle of turkeys" (Fouse 2002: 137).

Figure 5 Asymmetric Differentiation
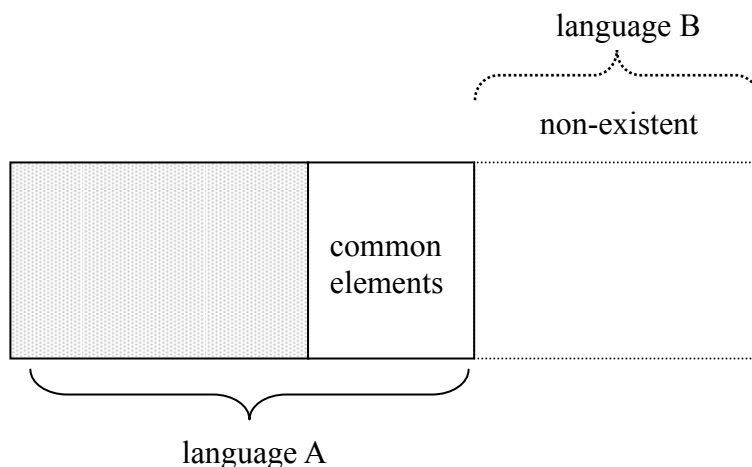
In fact many pidgins and creoles provide ample evidence for this. In general, pidgin and creole languages are often perceived as some thing "negative" or "less desirable" (Fouse 2002: 16). Tok Pisin (formerly called also New Guinea Pidgin or Melanesian Pidgin) which is the national language and lingua franca of Papua New Guinea, for instance, was regarded as "disgusting, debased corruption of English" which "[sounds] quite ridiculous" (Wurm 1985: 377). It was the same in the case of Yiddish. Yiddish is a Judaicized version of High German dialects (for sources or superstrate languages, see Goldsmith 1997: 30 and Weinsten 2001: 18-19, as well), formed in the German Lands around the ninth and tenth centuries by Jews (Wexler 1989: 142). 97% of the Russian Jews spoke the language in 1987 and it had daily newspapers, theatres and schools. And in the late 1930s, the Jewish autonomous province of Birobidzhan was established in the Far East. Despite the facts, Yiddish was subjected to denigration, scorn and criticism, with its lack of authenticity, beauty and autonomy (Peltz and Kiel 1985: 278-279).

Although these examples amply show that one of the two hypothetical languages in our consideration can easily be perceived not to be a real language, two question remains here. First, is the language perceived as non-existent completely out of the awareness? Secondly, if the language is perceived rather as defective or incomplete one lacking in crucial properties like "grammar," then is it better to include them into the category of "identification"? As we shall see shortly, it is very difficult to distinguish

above cases of "differentiation" from later examples of "identification." Perhaps, "exclusion" and "inclusion" may be more appropriate terms to be used than the terms "differentiation" and "identification." But we will leave this question open.

As to the first question, it must be confessed that the cases of complete unawareness as we discussed in connection with Figure 5 are rather rare as various citations suggest. The following example of so-called Black English (Vernacular) is perhaps the nearest case. Black English is, as many linguists agree, an independent language of its own, "a separate system closely related to standard English" (Labov 1970: 184). Therefore, the actual relationship between standard English and Black English is like that given in Figure 1, and the perception of the relationship is expected to be that of Figure 3 or 4 (perhaps 4). But the actual prevalent perception is rather like that in Figure 6. The dominant view, often referred to as "deficit theory," attributes the school failures of Black children to their intellectual deficits (Baratz 1970: 11-13). But the critics argued against the dominant view that what is deficit is the perception and understanding of Black English held by the dominant view rather than the intellectual capacity of Black children. In fact, the dominant view can be shown to be doubly deficit.

Figure 6 Deficit View

language B (Black English)

common
elements

language A (standard English)
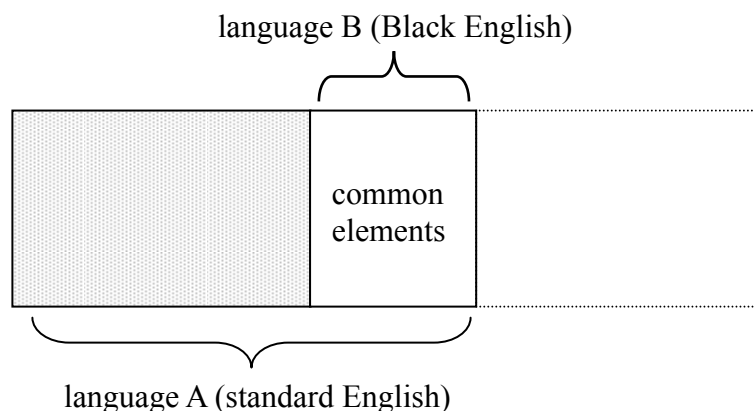
As is shown Figure 6, the two languages are distinguished in a very peculiar way in this case. First, language A is perceived as usual, as consisting of two parts: that is, the elements which are unique to it and the elements which are common to or shared

by the two. What is peculiar in this case is the perception of B. Language B is perceived as consisting only of one part: that is, the elements common to both. It is perceived not to have the elements of its own. The dominant view is completely unaware of the existence of the elements unique to B. If the two languages are compared, therefore, language B, Black English in this case, is seen as a language which lacks the unique elements of language A. Accordingly, language B, Black English, is a deficit language in the dominant perception. But, it is obvious from our previous examination that what is deficit is not Black English but the dominant view which cannot recognize the existence of the unique elements of Black English. Thus, the dominant view is doubly deficit, holding a partial view of the language and deeming the language to be deficit. The deficit view can be considered a clear case of exclusion. In this case, the language B (elements unique to Black English) is excluded from the perceived language.

3 Identification

Given the basic relation between two languages we have so far assumed, another type of perception, "identification," is quite possible. Let us now turn to the perception pattern of identification of the two languages in question. In identification, the two languages A and B are perceived to constitute one single language, and each are regarded as a dialect or variety of the common language. The typical case is shown in Figure 7.

Figure 7 Identification

language C
"Languages" A and B are perceived to constitute parts of Language C, and A and B are varieties (or dialects) of C.
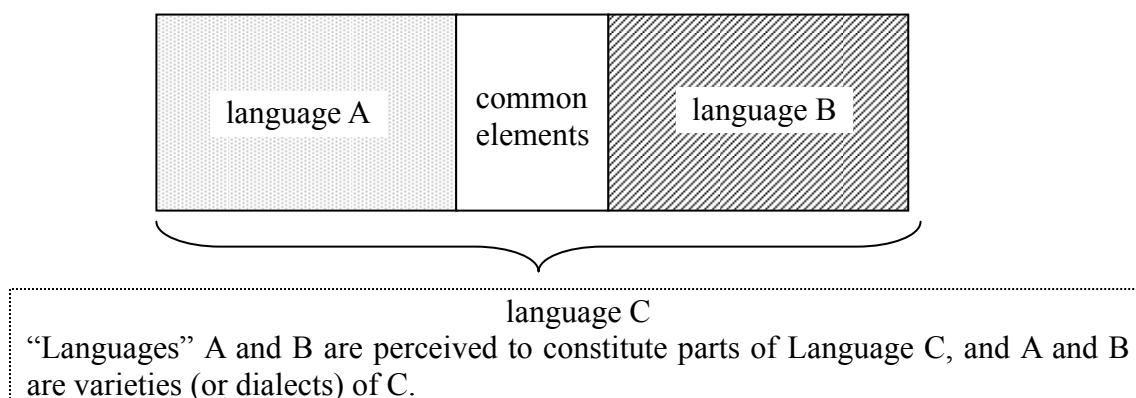
There are much more complicated cases as well. Djité refutes the alleged plethora of languages in Africa, and shows that "[t]he barriers to the use of African

languages as media of instruction or national and official languages are much lower than previously thought" (Djité 1988: 8). Djité demonstrated on the basis of the perception of native speakers that two languages in Côte D'Ivoire, Guéré and Wobé, which are officially and academically regarded as two distinct languages, are actually two varieties of one single language called the "Wɛ" language. She argues that these are actually one language with the common core (called language core or monolingual nucleus) and with peripheral differences (called multilingual satellites). Moreover, the "Wɛ" language cluster contain two more languages as is shown in Figure 8 (Djité 1988: 6-8). In this way, Djité convincingly shows that two (and actually more) languages perceived to be distinct in the official and academic perception can be just one language in another perception.

Figure 8 "Wϵ" Languages
source: Djité 1988: 8
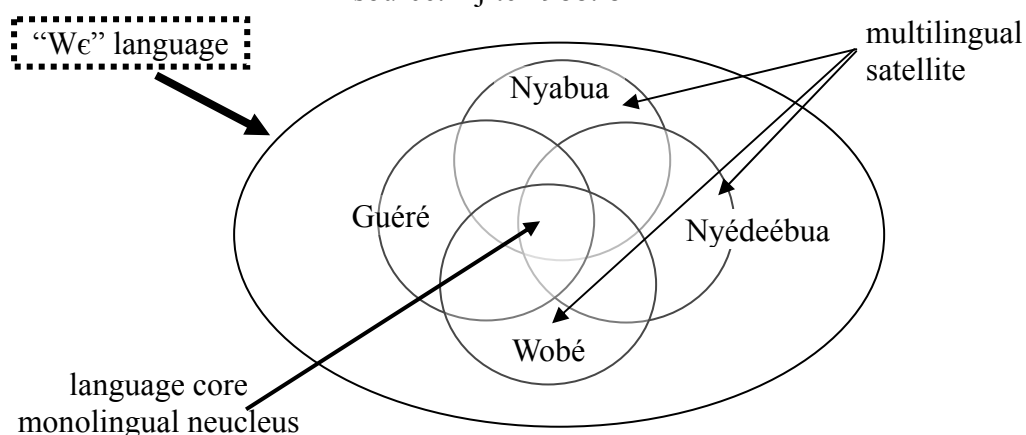
Even when our two languages (or dialects) seem far apart with very little in common, perceptual identification can still occur. One typical example is Mordvinian. Mordvinian is classified into the Volga-Finnic branch of Finno-Ugrian language family, and it is closely related to Finnish. It has two quite distinct dialects called Erzia and Moksha. Neither of the dialects has a word to designate the language itself. To refer to the language itself, therefore, one must use the phrase "Erzia and Moksha." But despite the difference in languages (or dialects) and other attributes, the Mordvinians have always been regarded as one and the same people, that is, Mordvinians. Modvinians are a people with one head but *with two mouths* (Kreindler 1985: 237-238). As this case of Mordvinian shows, a very great difference does not prevent the perceived identity of

languages, in some cases at least.

As a case of identification which is a little more complicated or a little closer to reality, we cannot ignore the issue of the standard language. A standard language is usually considered by linguists to be a superposed or umbrella variety of a language, usually based on, or created from, one of the language varieties (or languages in our word), as is shown in Figure 9. In this case, languages A and B together with the standard language are perceived to constituting one single language. And A and B are viewed as dialects of language C.

Figure 9 Superposed Standard Language

language C
"Languages" A and B are perceived to be dialects of Language C

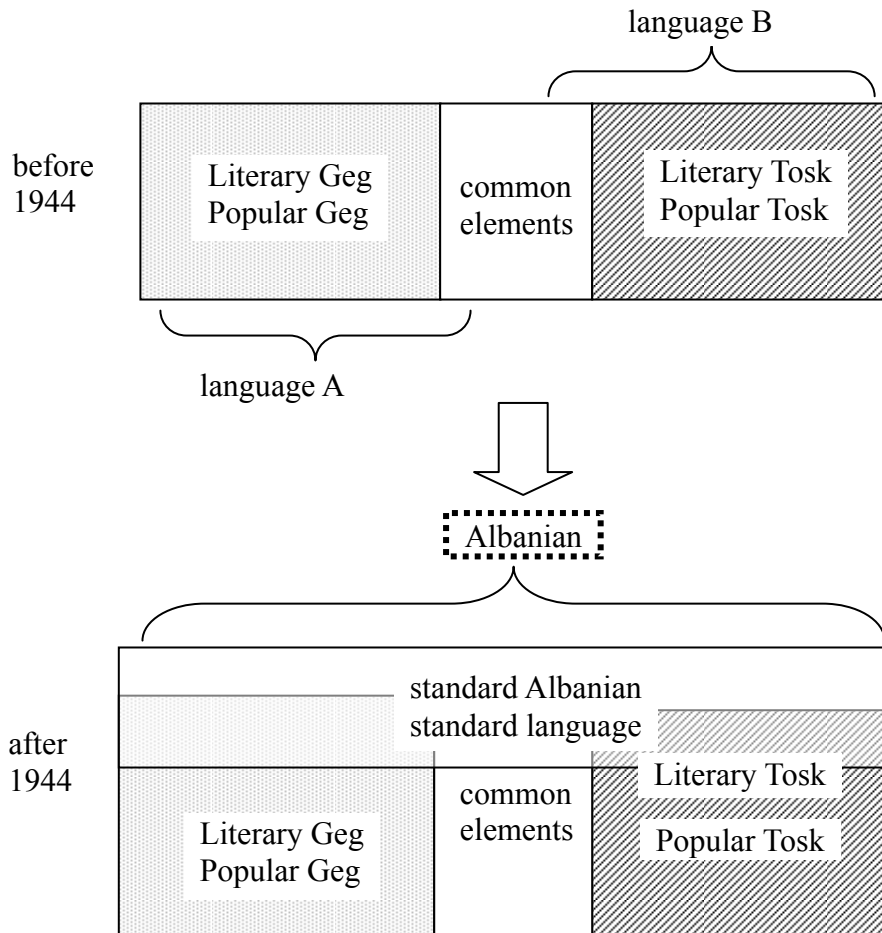| standard language		
language A	common elements	language B
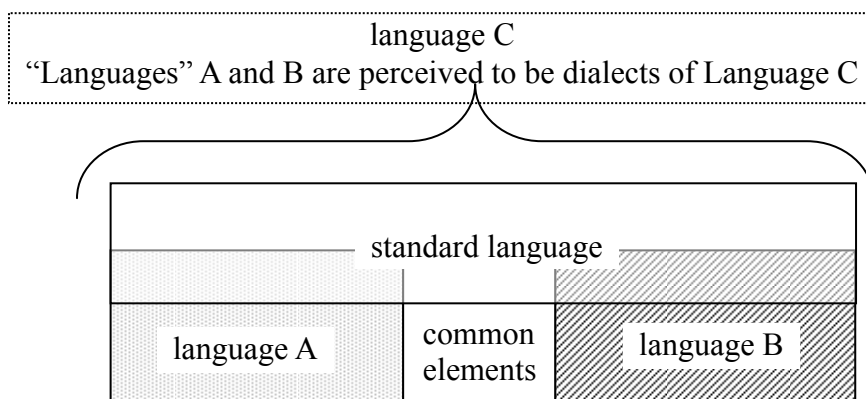
The Albanian case provides an example, though it is a little atypical. As is well known, there are two major dialects in Albanian; Geg in the north and Tosk in the south.

The lack of significant bidialectalism has been maintained despite the centuries-long contact of the two dialects (Byron 1976: 49). Thus before 1944, both had their own written and spoken varieties as is shown in Figure 10. But after the seizure of power by communists, the southern dialect Tosk was selected as a basis of the standard Albanian, because the Albanian communist movement originated in the south and majority of leaders were Tosk speakers coming from the south. The standard Alabanian was almost the same as the literary written Tosk at first. With the passage of time, however, the standard gradually become more and more removed from Tosk. Consequently, the standard Albanian today is rather a superposed variety in relation to other varieties as is shown in Figure 10, though it still shares most with literary Tosk

(Byron 1976: 59-74).

Figure 10 Albanian: Dialect and Standard
source: Byron (1976), 64

A similar but a little more complex case is found in Thailand. According to William Smalley, the Standard Thai language is a variety based mainly on Thaiklnag (dialect) and it is superimposed upon the four dialects: Kammüang, Thaiklang, Lao, Paktai (Smalley 1988: 247). To complicate the matter, of the four "dialects", Lao (or Isan) is actually the national language of Laos, and usually regarded as a separate language distinct from Thai. As far as Lao is concerned, the Thai case may also be considered a case of inclusion like Ukrainian which we will discuss below.

We have examined cases where the two (or sometimes more) languages in question are perceived rather similarly or equally than differently. But more often than

not, one of the languages is perceived as inferior to the other, as dialect or corrupted form of the superior other. Historical and contemporary examples abound.

In July 1863, Russian Minister of Internal Affairs, Petr Valuev, wrote in a letter that

> There has not been, there is not, and there can not be any kind of separate Little Russian [i.e., Ukrainian] language. … [T]heir *dialect*, spoken by the common people, is that *very same Russian language*, but *spoiled* by the influence of Poland; the all-Russian language is as comprehensible to the Little Russians as to the Great Russians and even more comprehensible tha[n] the so-called Ukrainian language that is currently being created for them by certain Little Russians and especially by Poles." (cited in Solchanyk 1985: 58, emphasis added).

This was the official stance of the Tsarist government toward Ukrainian. Accordingly, in order to forbid the use of Ukrainian and thereby suppress the rising Ukrainian national sentiment and movement, Valuev had to have recourse to a secret decree (Solchanyk 1985: 58), because the Ukrainian language was officially non-existent. Who can ban the use of something which has not existed, does not exist, and will not exist?

Figure 11 Inclusion
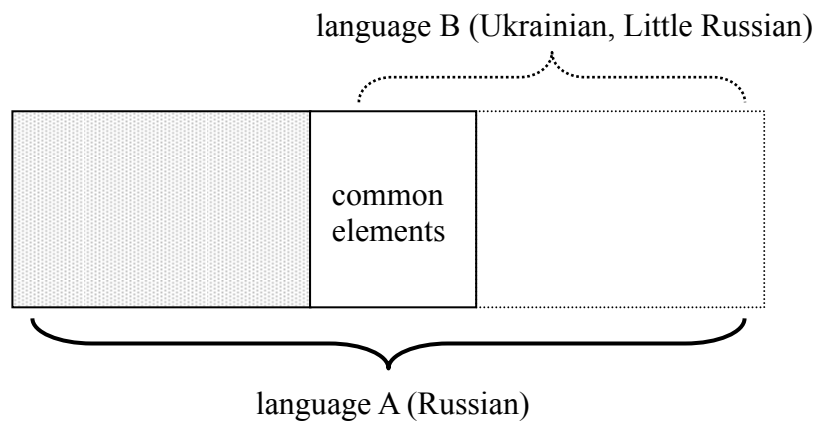
The Russian-Ukrainian case is a clear case of inclusion. The Ukrainian language is perceived to be a part (a dialect) of the Russian language

The case of Galician in Spain is a little more complicated. It is, as it were, a case of double inclusion. Though it is now accorded a status of a distinct language, it was formerly perceived to be a dialect of both Spanish and Portuguese (Green 1990 194, Parkinson 1990 250).

4 Conclusion

In this paper, we found various patterns of perception of the identity and difference between languages (or, to be precise, between language varieties), though they are not exhaustive. Languages seem, by their very nature, to allow a wide variety of perception patterns concerning their identity and difference. It is clear from a brief and simplified examination above that there is nothing objective in the nature of language which automatically determines the sameness or difference of two languages, as far as the languages in question are somehow related languages.

The possibility of different perceptions of one and the same reality of language relationship has an important political implication. First, between competing or opposing groups, there can be competing perceptions of the identity and difference between the languages in question. Secondly, the same is true even within a group. For example, in 1910s, the controversy between Pan-Turkism and linguistic particularism like Tatarization swept wide Turkic-language speaking areas from Central Asia to the Crimean Peninsula (Lazzerini 1985: 113-114).

In this paper, we dealt with patterns of perception, and did not deal with how they are produced, maintained, or transformed. This is an area to be investigated in future. In this connection, we should point out the importance of the issue of script, that is, by which character set (for example, Latin, Cyrillic, Devanagari etc) the language in question is written. It is because the script will give a very visible clue to the perception of the identity and difference of languages. Cases like Serbo-Croatian (Corbett 1990: 128) and Hindi-Urdu (Campbell 1991: 571, 1425, Schackle and Snell 1990: 7-11) where two varieties are written in different scripts reinforce the perception of difference.

These possibilities of different perceptions of language identity provide ethnonational leaders with handy political resources, and are exploited as such for political purpose, as many ethnonational movements have attested.

References

Aitkin, Adam J. (1990), "The Good Old Scots Tongue: Does Scots Have an Identity?" Einar Haugen et al (eds.), *Minority Languages Today*, Edinburgh: Edinburgh University Press, 72-90

Baratz, Joan C. (1970), "Teaching Reading in an Urban Negro School System," Williams (ed.), 11-24

Byron, Janet L. (1976), *Selection among Alternates in Language Standardization: The Case of Albanian*, The Hague: Mouton

Campbell, Geroge L. (1991), *A Compendium of the World's Languages*, 2 vols., London: Routledge

Chomsky, Noam (1975), *Reflections on Language*, New York: Pantheon Books

Comrie, Bernard (1990), "Introduction," Bernard Comrie (ed.), *The Major Languages of Eastern Europe*, London: Routledge, 1-19

Corbett, Greville (1990), "Serbo-Croat," Bernard Comrie (ed.), *The Major Languages of Eastern Europe*, London: Routledge, 125-143

Djité, Paulin G. (1988), "Correcting Errors in Language Classification: Monolingual Nuclei and Mulitilingual Satellites," *Language Problems and Language Plannign*, 12(1), 1-13

Fouse, Gary C. (2002), *The Story of Papiamentu: A Study in Slavery and Language*, Lanham: University Press of America

Goldsmith, Emanuel (1997), *Modern Yiddish Culture: The Story of the Yiddish Language Movement*, New York: Fordham University Press

Green, John N. (1990), "Romance Languages," Bernard Comrie (ed.), *The Major Languages of Western Europe*, London: Routledge, 193-199

Haugen, Einar (1990), "Danish, Norwegian and Swedish," Bernard Comrie (ed.), *The Major Languages of Western Europe*, London: Routledge, 147-169

Kreindler, Isabelle T. (1985), "The Mordvinian Langauges: A Survival Saga," Kreindler (ed.), 237-264

Kreindler, Isabelle T. (ed.) (1985), *Sociolinguistic Perspectives on Soviet National Languages: Their Past, Present and Future*, Berlin: Mouton de Gruyter

Labov, William (1970), "The Logic of Nonstandard English," Williams (ed.), 153-189

Lazzerini, Edward (1985), "Crimean Tatar: The Fate of a Severed Tongue," Kreindler (ed) (1985), pp.109-124

McClure, J. Derrick, A. J. Aitken and John Thomas Low (1980), *The Scots Language: Planning of Modern Usage*, Edinburgh: Ramsay Head Press

Parkinson, Stephen (1990), "Portuguese," Bernard Comrie (ed.), *The Major Languages of Western Europe*, London: Routledge, 250-268

Peltz, Rakhmiel and Mark W. Kiel (1985), "Di Yiddish-Imperye: The Dashed Hope for a Yiddish Cultural Empire in the Soviet Union," Kreindler (ed.) (1985), 277-309

Shackle, Christopher and Rupert Snell (1990), *Hindi and Urdu since 1800: A Common Reader*, London: School of Oriental and African Studies, University of London

Smalley, William A. (1988), "Thailand's Hierarchy of Multilingualism," *Language Sciences*, 10(2), 245-261

Solchanyk, Roman (1985), "Language Politics in the Ukraine," Kreindler (ed.), 57-105

Steinberg, Jonathan (1987), "The Historian and the Questione Della Lingua," Peter Burk and Roy Porter (eds.), *The Social History of Language*, Cambridge: Cambridge University Press, 198-209

Weinstein, Miriam (2001), *Yiddish: A Nation of Words*, South Royalton, VA: Steerforth

Wexler, Paul (1989), "Hieratic Components in Soviet Dictionaries of Yiddish, Dungan, and Belorussian," Björn H. Jernudd and Michael J. Shapiro (eds.), *The Politics of Language Purism*, Berlin: Mouton de Gruyter, 141-167

Williams, Frederick (ed.) (1970), *Language and Poverty: Perspectives on a Theme*, New York: Academic Press

Wurm, Stephen A. (1985), "The Status of New Guinea Pidgin (Neo-Melanesian) and Attitudes towards It," Nessa Wolfson and Joan Manes (eds.), *Language of Inequality*, Berlin: Mouton, 373-386