1 Introduction

In the late twentieth century and in the early twenty-first century, the prevalence of ethnic conflicts has aroused a great academic attention. Recently, however, a number of scholars have argued that allegedly ethnic conflicts themselves are not caused directly by ethnic difference or plurality (or multiethnicity) within a state\(^1\), though there are many arguments to the contrary\(^2\). But, in spite of the often alleged small relevance of ethnicity as a causal factor of so-called ethnic conflict, an important question is left still unanswered. The question we should ask now is why so many of violent internal conflicts today are ethnic conflicts, or more precisely, why so many internal conflicts are fought along ethnic fault lines or along ethnic boundaries. In this paper, we will explore one property of an “ethnic group” or ethnicity as a step toward the final answer to this question.

Though language is not the only factor or the most important factor which distinguishes ethnic groups, many scholars have so far argued that language is among the key factors in terms of which an ethnic group is distinguished from other group(s)\(^3\). And many authors admit that, in extreme cases, language is raised to the only symbol representing the entire relationship of the conflicting ethnic groups in question (Nelde 1987: 35, Ozolins 1996: 182). But these scholars have seldom gone beyond this categorical assertion, as if the relationship between a particular language and a particular ethnonational group is given and granted. The present paper is an attempt to push a little further by exploring the modalities of the relationships between language and an ethnonational group. It is because, as Uldis Ozolins once suggested, there are “a variety of possible relations between language and ethnicity in different historical context” (Ozolins 1996: 182). And we wish to answer why this is so.

In the next section, we will first prepare the ground on which to deal with the issue, stating premises and assumptions. In sections 3 and 4, we will make a brief theoretical examination of a property which functions as a boundary marker of ethnic
groups, and propose two conditions, internal homogeneity and external heterogeneity, which the boundary marker should satisfy. In section 5, on the examination of the nature of the relationship between two or more languages, we will show that the actual language relationship is such that both identification and distinction of the languages in question are possible. In other words, we will show that what matters is the perception or interpretation and not the reality as far as the identification and differentiation of languages are concerned. The next section, section 6, adds a few cases which show the relevance of perception in language identification and differentiation. In the last section, we conclude the essay by stating that it is perception and not reality that matters, and that this makes the association of a given language with a given ethnonational group very flexible.

2 Approach to the Relationship between Language and Ethnicity

As is well known, both language and ethnicity (ethnic or ethnonational group) are very elusive concepts. Volumes of books will be required for the definition of these terms. As a first approximation, therefore, we will be obliged to start with preparing the ground for our discussion.

We have first to solve, or clarify at least, the issue of terminology for one of our variables. It is because there are so many, and perhaps too many and confusing, terms for ethnicity or ethnic group. For example, besides the term “ethnic,” we find such terms as “communal,” “national,” “nationality,” “mini-national,” “subnational,” “ethnonational,” “ethnopolitical,” “ethnoregional,” “ethnoterritorial” and so on. In connection with ethnic group or ethnicity, the term “nation” is very frequently used in a similar sense, though some authors distinguish them strictly. In the present paper, we will employ the term “ethnonational group” or “ethnicity” rather than other terms in order to cover both a “nation” and an “ethnic group,” or to be able to explore the issue both at the state and sub-state levels. Then, even when a “nation” and an “ethnic group” can be, or should be, distinguished conceptually, the arguments in this paper mostly, if not completely, apply to both of them. In addition, we will not use such subcategories or subdivisions as “nationality,” “ethnic minority,” “indigenous peoples,” “tribe” or “clan” and so on, either. Though these and many other terms may have their own usefulness in a certain circumstance, we will stick to the use of the general term “ethnonational
Secondly, we have to limit our attention rather narrowly upon a certain facet of the overall relationship between language and an ethnonational group. In this paper, we will focus upon one aspect of the relationship, that is, the issue whether and how language defines an ethnonational group. To put it in a little different way, our focus is language as a boundary marker of ethnonational groups. Such a narrow focus of course leaves many important issues untouched. Let us point out two of them.

One is the omission of the cases where an ethnonational group seems to define language, but not *vice versa*. According to Earnest Gellner, the modern state has created a language community characterized by a common written language or “high culture” (Gellner 1983: 33-34). As William Safran points out, the characterizing relations between language and an ethnic group are bidirectional (Safran 1999: 92). Language and ethnicity constitute, and are constituted by, the other. However, we focus upon language as a distinguishing marker of an ethnonational group. In addition, we must admit of another serious omission. Identity and interest constitute two important dimensions of the language-ethnicity relation. Our focus on language as an ethnicity marker leave the latter out of consideration as well, despite of the increasing importance of the latter (Glazer and Moynihan 1975: 7, 12-13, Inglehart and Woodward 1967: 32).

Thirdly, when we deal with the relationship between variables, language and an ethnonational group, it is theoretically convenient, and perhaps more rewarding, to keep one variable constant. Therefore, we will assume in our discussion that an ethnonational group is given and fixed, except in the cases where language seems to cause the change of an ethnic group, though it does not mean that we believe in the primordial fixity of the group.

Fourthly, it is generally argued that in order to be an ethnonational group, or any human group for that matter, a human collectivity should share both some objective properties and the sense of belonging to a community. People may share some properties like language or customs without this shared sense of belonging. But they don’t constitute an ethnonational group without the shared sense of belonging. Even from our limited perspective of language as an ethnicity marker, therefore, as Anthony Smith warns (Smith 1986: 211), we should address both the objective aspects and the subjective aspect of the sense of belonging. But here we will not go into the discussion.
of the sense of belonging in this paper. We will only deal with language as an objective marker of an ethnonational group. But this apparently objective property involves a great deal of subjective dimensions as we will see below.

3 Shared Objective Properties

It is usually argued that an ethnonational group is a group of people sharing a certain objective properties. Objective properties or attributes here include ancestors, history, territory (place of “ancestral” habitation), economic life, language, religion, culture, behavior pattern, physical characteristics (like skin color) and so on. They are largely given to an individual or cannot be changed easily by an individual, and they generally inherited through generations.

If these properties are to define an ethnonational group, that is, to distinguish an ethnonational group from another, they should be shared by its members and inherited from one generation to another. Most of these properties are visible objective properties. They provide an objective foundation for the existence of an ethnonational group. To repeat, an ethnonational group is a collectivity of people sharing these properties.

There is an interesting question concerning the sharing of objective properties. Should an ethnonational group share all or most of the properties or attributes mentioned above, or how many of them should it share? Since we focus only on language, we will simplify this question into: whether an ethnonational group should share a language or not. In later sections, we will try to answer this question.

On the other hand, the sharing of one or more properties by two or more groups does not guarantee that they belong to, or constitute, one single group. Shared objective characteristics are, as it were, raw material for the criteria for defining ethnic group, rather than the criteria themselves (Rothchild 1981: 95). It is also true of language. Sharing of the same language does not guarantee that the users belong to the same ethnonational group (Das Gupta 1975: 470, Safran 1999: 81).

From this, we can conclude that objective properties shared by an ethnonational group should be those which can be used to distinguish the group from others. Let us call such properties “distinctive.” Distinctive properties are those properties which function as boundary markers between ethnonational groups. Which of the shared
objective properties plays the role of distinctive feature depends upon the historical conditions of the ethnonational group in question. The conflict of two major ethnonational groups in Belgium has been well known. Because they are both Catholics and hence share the religion, language (Dutch and French) and territory or region (North and South) became the distinctive properties distinguishing the two groups. Similarly, in the independent India, the partition deprived the function of an ethnicity marker of the religion, and raised language to an ethnicity marker, leading soon to the creation of the “linguistic states” beginning with the case of Andhra Pradesh (Inglehart and Woodward 1967: 32). On the contrary, religion and religious denomination or sect are distinctive properties in Northern Ireland and Lebanon. In sum, what matters in distinguishing ethnic groups is the distinctive properties, and, as an ethnicity boundary marker, language should be a distinctive property whether separately or combined with others.

Here we should make a reservation on the scope of distinction or distinctive property. A distinctive property or a set of distinctive properties need not distinguish an ethnonational group from all the other ethnonational groups. It can be easily understood when we consider the situations in which ethnic belonging matters. Some scholars say that an ethnonational group becomes worth the name only when it is in contact with other groups. In other words, an ethnonational group becomes truly an ethnic group when and only when the group needs to be distinguished from another group. Accordingly, a distinctive feature or a set of them have only to distinguish the group in question from those who are in contact with it in some way or other. Thus, if an ethnonational group is not in contact with any other group, it has no need to distinguish itself from others theoretically at least. Consequently, it has no need to have a name to distinguish itself form others. Thus, the fact that many indigenous peoples of the world such as Ainu in Japan, Maoris in New Zealand, and many Siberian ethnic minorities like Evenki have used self-names meaning simple “human being” is the other side of the same coin.

A distinctive property or a set of them have only to distinguish the ethnonational group from a certain particular group or groups, but not from all the other groups. The objective properties are needed only to be distinctive, but not to be unique, that is, proper to the ethnic group alone or distinguish it from all the others. The French language is a distinctive property to the French Canadians or to Quebecers (or
Quebecois) which distinguishes them from the English-speaking majority (Anglophones) in Canada, but the French language cannot distinguish them from the French people in France, because there is no need for the French Canadians to distinguish them from the French people in France. If it was necessary to distinguish French Canadians and the French in France by language, the distinction between the standard French and its Quebec dialect called “joual” (Weinstein 1989: 54, 57-58) would be resorted to.

In sum, a distinctive property should be viewed as property which can distinguish the ethnonational group in question from certain other(s), but not from all the others. It is true, however, that in many cases a distinctive property or a set of them tend to be perceived as unique to the ethnonational group in question, that is, as the property which separates the group from all the others. From this perspective, it is quite natural that, in order to consecrate the uniqueness of the group, some groups should attempt to establish their uniqueness in name, even when they have actually no unique or particular property at all. Indeed, it is commonplace. To call the Malay language Indonesian or Malaysian is a typical example.

4 Internal Homogeneity and External Heterogeneity: Conditions for a Distinctive Property

So far, we have used the term “share” without any definition. In this section, let us examine what “share” means, with special reference to language.

Suppose X represents a distinctive property which distinguishes the ethnonational group in question from certain others. In order to be a distinctive property, X must satisfy the following two conditions.

1. X is shared by numerous enough members of the ethnonational group (homogeneity or similarity proposition)

2. X does not apply to (is not shared by) most of the members of certain other ethnonational group(s) (heterogeneity or difference proposition)

The first proposition requires that the ethnic group be homogeneous internally as far as X is concerned. The second requires that the two groups to be compared be heterogeneous to each other as far as X is concerned. Combined, the proposition
requires internal homogeneity and external heterogeneity of X. In other words, X must be shared by the members of the ethnic group in question but must not be shared by members of certain other ethnic group(s). When we examine language from the perspective of an ethnonational boundary marker, it may first appear that we have only to replace X in the propositions above by language. But a difficult question arises at once. Should the language which allegedly distinguishes an ethnonational group from some others be both spoken and written by the majority of the group members, or is it sufficient for it to be spoken or understood? We will take up this issue again later in more detail.

The relevance of the two conditions can be observed in historical and contemporary language conflicts, especially in those I called “nested conflict.” In nested conflicts, an ethnonational group seeks both to differentiate itself in terms of language from others and to homogenize the language within, as in the 19th Kingdom of Hungary and in the Post-Soviet Moldova (Matsuo 1999: 96-99). 5) It can be said that, in these and other cases, the two conditions are turned into a kind of norm.

The two conditions are not only politically important but also in theoretical consideration. When these conditions, internal homogeneity and external heterogeneity, are not met, many previous studies on ethnicity have erroneously drawn either of the two conclusions. On the one hand, the (set of) distinctive property(ies) may be concluded not to be the defining characteristic of the ethnonational group in question, because they are not shared by the members or they are shared by the members of certain other group(s). On the other, it may be concluded that the group is not a distinct and separate ethnonational group because the group does not satisfy the conditions necessary distinctive properties. In the latter case, the distinctive property, X in our case, is treated as an absolute criterion in that every ethnonational group should share the property X. Both of these conclusions are erroneous in most of the cases.

The most serious error in these arguments is that they have fallen into the fallacy of thinking of the issues only in reality and completely ignore the issue of perception. When we consider the relationship of a distinctive property, especially of language, with an ethnonational group, we should consider them both in reality and in perception. The latter perspective, that is, the perspective of perception, is indeed essential.
Though we will examine the issue in much more detail later with reference to language, let us now consider, as an illustration of the importance of perception, the problem of the ancestry or the origin of an ethnonational group. Common ancestry is regarded as one of the objective bases of the group. If one goes back in history, however, it is often the case that the genealogy of an ethnonational group disappear somewhere in the dark past of history at least in the strict scientific sense of the word. Moreover, it is difficult to find a genuinely “pure” ethnonational group without any mixture with other groups. In this sense, if we stick to the common ancestry as the basis or distinctive property of an ethnonational group, there will be no such thing as ethnonational group in the objectively strict sense.

But as far as ethnicity is concerned, what is wrong is not the concept of common ancestry, but the denial of the status of ethnonational group on the ground that it has no common ancestors in the objective historical sense. In other words, the error does not lie in the fact that there is no group which shares a single common ancestry in the strict sense, but it lies rather in the premise that an ethnonational group should have a single common ancestry as an objective historical fact. As Walker Connor insightfully points out, an ethnic group is the greatest human group characterized by sharing the myth of common ancestry, but it matters little whether the myth corresponds to the historical fact (Connor 1987: 211).

Generally speaking, what is wrong is the thought that distinctive properties are objective (or historical) facts and realities. To put it in a little different way, what matters is whether members of the group in question (and non-members as well) believe in the common ancestry or not (Connor 1987: 205-206). To be sure, it is clearly an overstatement to say that no visible or objective characteristic is essential to the existence of an ethnonational group (Connor 1987: 202). But, an ethnonational group need not actually share a common ancestor(s). And, generally speaking, an ethnonational group need not actually share distinctive features. The sufficient condition is that the distinctive property, the common ancestry in this case, is perceived or believed to be shared by members of the group in question. Thus, what matters is not the actual sharing, but the perceived sharing of properties. It is especially relevant to language as an ethnicity boundary marker as we see shortly.

From these observations, we can say at this stage of our investigation that a (set
of) distinctive properties which distinguish an ethnonational group from certain other group(s) need not be actually shared, but need only be perceived or believed to be shared by members of the group in question. This does not mean, of course, that distinctive features cannot or should not be actually shared by members. Such perception of sharing properties can function as evidence of the existence of an ethnic group, can influence behaviors of its members, and can have an important political effect. Thus, the perceived shared properties like common ancestry can function as a myth.

5 Language Relationship: Reality and Perception

In the previous section, we proposed two conditions to be satisfied when a language (or any other property) distinguishes an ethnonational group from others. The language should be shared by the members of the group, but not shared by members of other group(s). The conditions involve three major components or one predicate and two arguments: sharing of a language, an ethnonational group, and a specific language. In order to clarify the role of language as an ethnonational group marker, the ethnonational group must be identified, and separated from other ethnonational group, and the language which distinguish the group from others must be identified and separated from other languages. As we said above, we here assume that the identification of an ethnonational group is given. Therefore we will now examine how a language is identified.

For a language to function as a boundary marker of ethnicity, the language should be distinguished from other languages. It is, however, a very difficult task even for linguists to decide whether two related languages (or language varieties) are the same language or two different languages, 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).

I once tried, in Matsuo (2005), to explain why such a difficult task of differentiation (or identification) of languages is not only possible but also has been often done “on political and social grounds.” The relationship of two languages (or
dialects) can be schematically represented as Figure 1.

![Figure 1 Basic Relationship](image)

On the one hand, the two languages in question differ from each other in grammatical rules (syntax), in lexical elements (vocabulary), or in sound system (pronunciation) or in some combination of them, or in all of them. On the other, however, there are always common elements (even if exclude loan words) which both share. Though it is not very accurate, the description will be sufficient from a purely linguistic point of view. Actually, more than two languages (or language varieties) are often involved as in the cases of dialect continua and creole continua (Comrie 1990: 3, 17, Friedman 1993:160, Kalogjera 1985: 94). Even for national language, it is possible to say that, as to Nordic languages, “Norwegian is Danish spoken in Swedish” (Haugen 1990: 151).

But, contrary to linguists’ assertion that “[w]riting is not language, but merely a way of recording language by means of visible marks” (Bloomfield: 1969: 21), writing system and set of letters (script) should also be considered to be an essential component of language, as researchers has found (Calvet 1998: 153, Schieffelin and Doucet 1998: 285). In fact, many conflicts have occurred over the writing system. To mention a few, the mid-19th Galicia under the Polish rule (Subtelny 1988: 513-514), East Pakistan (now Bangladesh) before independence (Musa 1996: 75-76), Moldavian Soviet Republic (Eyal 1990: 126-127), Latinization of script for Turkic languages in the Soviet Union (Lazzerini 1985: 116-117) are cases in point. Moreover, even if it does not lead directly to violence, the writing system has been employed to distinguish ethnonational
groups, whether singly or in combination with other elements. Standard Serbian and Croatian in the former Yugoslavia are distinguished by Cyrillic and Latin scripts (Corbett 1990: 128), and Hindi and Punjabi are distinguished by Devanagari and Grumukhi scripts (Dale 1980: 10, Dua 1996). The relationship between two languages (or language varieties like dialects) given in Figure 1 above is assumed to represent the reality. Of this one single reality, as was shown in Matsuo (2005), a variety of interpretations are possible. The real language relationship is such that it permits us different interpretations. As in other social phenomena, one fact permits many interpretations. Thus, while some can identify two languages as one, viewing the two as dialects of one single language, others can differentiate them, viewing them as two separate languages. Though numerous examples can be adduced for both, let it suffice to give one or two examples for each.

First, in spite of vast geographical differences, Chinese and Arabic are each regarded as one single language. English may be another example. This kind of identification is not limited to such “big” languages spoken in wide areas and by vast number of people. For example, Mordvinians have been viewed as a people with one head but with two mouths because they have two dialects mutually nearly unintelligible; Erzia and Moksha (Kreindler 1985: 237-238).

In the late nineteenth century, Russian authorities regarded Ukrainian (or Little Russian) as a dialect of Russian though it was contaminated by Polish influence, while Ukrainian nationalists declared it to be a separate language (Solchanyk 1985: 58). And Belorussian was variously regarded as a dialect of Polish, Russian or Ukrainian (Wexler 1985: 38). Perception of Galician in the Iberian Peninsula (Parkinson 1990 250) provides us with another example.

We can also find many cases of perceptual differentiation. A great similarity (or mutual intelligibility) does not necessarily prevent differentiation of two languages (two language varieties). Some argue that Scots (which is usually called Scottish English and, as the name suggests, perceived to be a geographical dialect of English) is a language of its own distinct from English (Aitkin 1990: 76). Perhaps, India has been a kind of laboratory for this kind of perception because identification of a separate language has been inseparably connected with ethnonational autonomy in the form of the “linguistic state” (Chaklader 1990: 130-131, 207-208).
6 Perceived Sharing of a Language

In the previous section, we showed that languages (or language varieties like dialects) can be identified or differentiated in perception largely irrespective of their actual closeness or distance. In this section, we will examine the predicate of our two conditions, that is, sharing.

It is theoretically possible to find out extent to which a given language is shared by members of an ethnonational group (in whatever way “sharing” may be defined). And, in fact, many such attempts have been made in the form of a census containing the question of mother tongue, first language etc. In this case again, however, it is not the reality but perception that matters. In some cases at least, a language need not be spoken or understood by the majority of a given ethnonational group.

The Kingdom of Hungary was under the Habsburg rule in the 18th century. In 1784, the emperor, Joseph II, declared that German should be the official language of administration and education throughout the empire (Fichtner 1997: 24). The “Germanization” declaration caused a violent reaction of the Hungarians, especially the nobility (Barany 1971: 262-263). In opposing the imperial policy of Germanization, the Hungarian nobility embraced the Magyar (Hungarian) language as their common language, though the Hungarian language seemed to be obsolescent among the nobility (Inglehart and Woodward 1967: 34). As in this case, the language sharing is not an objective necessity, but it is rather a(n) (inter)subjective necessary condition. In other words, it is sufficient that a language (Hungarian in this case) is perceived (or believed or imagined) to be shared by the members. In this sense, an ethnonational group can be an imagined community (Anderson 1983: 15) based on a language.

The Irish case is perhaps the most illuminating example of the importance of perception of language sharing. Let us here look at some statistics on the Irish language. Gaelic (Irish)-speakers were less than 20 % of the population before independence in 1922 (Hindley 1990: 15, 23). It cannot be said that the language is actually shared by the members. Nevertheless, the Gaelic language is a distinctive feature of the Irish people in perception and in a symbolic sense. Gaelic is perceived to be the language of the Irish people. It is for this reason that the Irish government declared Gaelic as the national language, made efforts for its promotion and taught it in schools.

In the independence movement in the late nineteenth and early twentieth
centuries as well, the Gaelic language was a symbol of the movement. Thus, the Gaelic League, which aimed at the revival and promotion of the Gaelic language, played a key role in the independence movement. In this case, the Gaelic language was not actually shared by the majority but played a great role in the Irish independence movement as a symbol of the group. The case of Ireland typically shows that what matters is not fact or reality of sharing, but perceived (or imagined) sharing.

In sum, a language needs not be shared actually by the majority of the members of an ethnonational group, it is sufficient, in some cases at least, that the language is perceived or believed or fostered by the members.

7 Conclusion

In the present paper, we first proposed two conditions which should be satisfied when a language functions as an ethnonational marker. Then we showed that the satisfaction of the conditions does not depend upon the reality but largely and doubly depends upon the perception. On the one hand, languages can be identified or differentiated rather freely, if not arbitrarily, irrespective of the difference or similarity. On the other, the sharing of a given language by members is sometimes a matter of perception rather than a matter of fact.

Therefore, the association of a language with an ethnonational group (even when it is held constant as we did) can be very flexible. And, therefore, language provides a very flexible and convenient criterion by which an ethnonational group is identified.

In this paper, we did not deal with the malleability of ethnonational identification. If we take the variation of ethnonational identification, the possibility of different ethnonational identification can expand greatly. Nor did we deal with policies or language movements which attempt to match the reality to the perception, for example, by imposing a dominant or majority language upon the minority groups. If we take into consideration such an attempt at the substantiation of the two conditions, it also expand the possible relation of language to an ethnonational group. And these issues should be among our future research agenda.

Notes

2 Though we don’t go into the details, there are many intermediate arguments. For example, ethnic polarization into two groups or the predation of a majority group may heighten the risk of internal conflict while ethnic fractionalization itself does not (Collier and Hoeffler 2000: 26, Collier and Hoeffler 2003: 537-538).

3 For example, see Das Gupta 1975: 470.

4 For such an argument, see Fishman (1999), 444-5, 449

5 In many other cases, either homogenization or differentiation effort is predominant. For examples, see Matsuo 1999, 92-97

6 Thus an author can say: “[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” (Gianrenzo Clivio cited in Steinberg 1987: 199).

7 Even in an extreme case of genetically unrelated languages, they share properties common to all human languages. Cf. Chomsky 1975, 29,

8 For examples, see Matsuo (2005), especially 192-202.

References


Chaklader, Snehamoy (1990), Sociolinguistics: A Guide to Language Problems in India, New Delhi: Mittal

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


Comrie, Bernard (ed.) (1990a), The Major Languages of Western Europe, London: Routledge


Fichtner, Paula Sutter (1997), The Habsburg Empire: From Dynasticism to Multinationalism, Malabar, FL: Krieger


Glazer, Nathan and Daniel P. Moynihan (1975), "Introduction," Glazer and Moynihan (eds.), 1-26


Haugen, Einar (1990), "Danish, Norwegian and Swedish," Comrie (ed.) (1990a), 147-169


Inglehart, Ronald F.and Margaret Woodward (1967) "Language Conflicts and Political Community," Comparative Studies in Society and History, 10(1), 27-45


Matsuo, Masatsugu (2005), "One Language or Two? Real and Perceived Identification and Differentiation of Language," Hiroshima Peace Science, 27, 189-203


Safran, William (1999), "Nationalism," Fishman (ed.), 77-93


