On Dialect Continuum

Pavić, Stana

Sveučilište J.J. Strossmayera u Osijeku

Filozofski fakultet

Preddiplomski studij engleskog jezika i književnosti i njemačkog jezika i književnosti

Stana Pavić

On Dialect Continuum

Završni rad

doc. dr. sc. Gabrijela Buljan

Osijek, 2011.

1. Summary and key-words

This paper deals with the issue of dialect continuum, which is a range of dialects spoken in some geographical area that are only slightly different between neighbouring areas. A dialect is not superior to another one, and certain dialects are considered languages mostly because of historical, political, and geographical reasons.

We cannot discuss dialect continuum without mentioning the cumulative, Levenshtein, geographic, and phonological distances. Those are "tools" used to study and explore dialect. Lines marking the boundaries between two regions which differ with respect to some linguistic feature are called isoglosses, and they can cause a number of problems because they are not always "perfect", they do not always coincide. Certain political and cultural factors created boundaries between different varieties of a dialect because a shared language is seen as something very important in the shared culture and economy, and a distinct language is important in demarcation one state of another.

Some of the most important dialect continua in Europe are the Western Romance, the West Germanic, and the North Slavic dialect continuum. Also the 27 Dutch dialects that lie on a straight line are a good example of how a dialect continuum looks like. With Levenshtein distances the linguistic distance between all these dialects is calculated. Another interesting, and very good example of dialect continuum is the dialect continuum of Ancient Greek that looks like this: West Greek ←→ Boeotian ←→ Thessalian ←(Lesbian → Ionic , A.-Cypr.)
where the two neighbouring ones share common linguistic features.

Key words: language, dialect, dialect continuum, Levenshtein distance, isogloss

Sažetak i ključne riječi

Ovaj završni rad bavi se pitanjima o narječnom kontinuitetu, a to je niz narječja koji se govore na nekom zamljopisnom području, a samo malo su drugačiji od narječja susjednih područja. Narječje nije superiorno u odnosu na neko drugo narječje, a određena narječja smatraju se jezicima ponajviše zbog povijesnih, političkih i zemljopisnih razloga.

Ne možemo razgovarati o narječnom kontinuitetu, a da ne spomenemo kumulativnu, Levnshtein, zamljopisnu i fonološku udaljenost. To su "alati" koji se koriste za proučavanje i istraživanje narječja. Linije koje obilježavaju granicu između dviju regija koje se međusobno razlikuju s obzirom na neke jezične značajke se nazivaju izoglose i one mogu uzrokovati niz problema jer nisu uvijek "savršene", tj. ne podudaraju se uvijek. Određeni politički i kulturni čimbenici stvorili su granice između različitih vrsta narječja, jer zajednički jezik se smatrao kao nešto vrlo važno u zajedničkoj kulturi i gospodarstvu, a različiti jezik je prema tome važan u odvajanu jedne državi od druge.

Neki od najvažnijih narječnih kontinuiteta u Europi su zapadnoromanski, zapadnogermanski i slavenski narječni kontinuitet. 27 nizozemskih narječja koji se nalaze na ravnoj liniji također su dobar primjer kako narječni kontinuitet treba izgledati. Uz pomoć Levenshtein udaljenosti se izračuna jezična udaljednost između tih narječja. Još jedan zanimljiv i vrlo dobar primjer narječnog kontinuiteta je narječni kontinuitet stare Grčke koji izgleda ovako: Zapadnogrčki ←→ Beocijanski ←→ Tesalijski ←(Lesbijanski

A.-Cipr.) → Jonski , gdje svaka dva susjedna dijele zajedničke značajke.

Ključne riječi: jezik, narječje, narječni kontinuitet, Levenshtein udaljenost, izoglosa

2.Introduction

The aim of this paper is to explain the term dialect continuum, some basic problems and principles in the interpretation of dialect data, and also the exploitation of linguistic situations to a variety of ideological and political purposes. Some linguistic problems concerning the implication of segmentation and separation, the impact of methodical orientation (synchrony and diachrony) on the interpretation on the linguistic image, and certain socio-political implications will also be explained.

The paper will start with the question what a dialect actually is and how it can be distinguished from the concept of 'language'. What are the reasons and factors that make a linguistic variety a *language*? The following section will deal with the definition of dialect continuum because it is a crucial thing to know this before discussing it in more detail. It will also explain the function of some basic instruments such as the Levenshtein, cumulative, geographical and phonological distances, all used when talking about dialect continuum, and the principles of how to make boundaries (isoglosses) between certain regions which differ with respect to some linguistic features and what the function of isoglosses is. The next two sections will deal with the concept of the geographical and social dialect continuum. After naming some most important geographical dialect continua in Europe, it will be shown that the reason for the creation of boundaries between different varieties of a dialect are not always linguistic, but socio-political. The next two sections provide some very interesting good examples of dialect continua. In the first one we will illustrate the example of a Dutch dialect continuum and use it to explain the usage of Levenshtein distance. The second example is the dialect continuum of Ancient Greek where the basic principles of dialect continuum will be illustrated, including the role of synchrony and diachrony when discussing dialect issues.

3. Dialect

This section will be concerned with the definition of dialect since an understanding of this concept is crucial for the topic we are discussing in this paper.

One problem in the definition of dialect is that the scientific idea of the term runs parallel to the lay conceptions of dialect. In the opinion of non-linguists, a dialect carries negative connotations; it is considered corrupt and is associated with rural values, as can be gleaned from the following quote:

> In common usage, of course, a dialect is a substandard, low-status, often rustic form of language, generally associated with the peasantry, the working class, or other groups lacking in prestige, but when talking about a dialect in linguistic terms, after exploring languages it can be said that all speakers are speakers of at least one dialect, and this dialect is the standard language, because standard English, for example, is just as much a dialect as any other form of English, so that we cannot say that some dialects are superior to another ones (Chambers and Trudgill 2004: 3).

The lay conception notwithstanding, it is difficult, even for linguists, to decide what a language and what a dialect is. One way of looking at this has often been to say that "a language is a collection of mutually intelligible dialects", but this is also not always true, because there are some languages that are mutually intelligible (Norwegian, Swedish, Danish), the speakers of those three language can communicate with each other, but they are considered to be three different languages (Chambers and Trudgill 2004: 3). Also German is considered a single language, but there are some types of German which are not mutually intelligible to speakers of other type.

3.1. Autonomy and Heteronomy

Another pair of concepts that are often invoked to explain the difference between dialects and languages is heteronomy and autonomy. Heteronomy suggests some kind of dependence rather than independence. This means that when there is a dialect of a language, that dialect is heteronomous with respect to his standard language. For example, speakers of the German dialects consider that they are speaking German, that they write and read in German, and that any changes in the dialect will be affected towards standard German. This means that German dialects, in a way, depend on the standard German language. Autonomy, in turn, would imply lack of such dependence.

Autonomy and heteronomy, however, are more a result of cultural or political than linguistic factors and changes. A good example is what is now southern Sweden, which was a part of Denmark in the 17th century, and the dialects spoken on that part of the Scandinavian dialect continuum were considered to be dialects of Danish (Chambers and Trudgill 2004: 9-10). However, after the war the territory became a part of Sweden. Only forty years later the same dialect were now considered dialects of Swedish. Here it can be seen that the dialects did not change themselves, but for political reasons they became heteronomous with respect to the new, standard language.

Some heteronomous varieties can also achieve autonomy because of political reasons and political developments. Chamber and Trudgill say that in this case 'new' languages can develop. An example for this is Norway. Earlier, in the 19th century the standard language there was Danish, and when Norway became an independent country, Norwegian was developed as an autonomous standard language. There are also some cases of semi-autonomy and "double or shared autonomy" (Chambers and Trudgill 2004: 11). Earlier, North American English saw British English as its 'standard' language, but today British, American and Canadian English (different forms of English) are 'on the same length'.

At the end there is a conclusion that certain varieties are considered single languages not only because of linguistic reasons, but because of historical, political, and geographical reasons , etc. In that case, one would rather speak of standard languages, which serve as vehicles of communication and act as symbols of, typically, national identity. "It is of course relevant that all three Scandinavian languages have distinct, codified, standardised forms, with their own orthographies, grammar books, and literatures; that they correspond to three separate nation states; and that their speakers consider that they speak different languages" (Chambers and Trudgill 2004: 5). Keeping these things in mind, we may say that it is difficult to establish an entirely linguistic definition of the term 'language' (although the linguistic branches of dialectology, comparative linguistics and typology do operate with their own definition of this term). But once we include socio-political criteria in its definition, things become more complex. For our purposes, we shall not insist on a technical definition for the term 'language', but will proceed with the following view of the concept 'dialect': "varieties which are grammatically, lexically and phonologically different from other varieties" (Chambers and Trudgill 2004: 5). Since dialects rarely stand in isolation, but pattern closely together in physical space, we will address in the following section the idea of dialect continuum.

4. Definition of Dialect Continuum

This section will deal with the definition of dialect continuum, because it is crucial to know what a dialect continuum is, before discussing it in more detail.

> If we travel from village to village, in a particular direction, we notice linguistic differences which distinguish one village from another. The further we get from our starting point, the larger the differences will become. At no point is there a complete break such that geographically adjacent dialects are not mutually intelligible (speakers of these dialects can't readily understand and communicate with one another), but the cumulative effect of the linguistic differences will be such that the greater the geographical separation, the greater the difficulty of comprehension. This type of situation is known as a geographical dialect continuum (Chambers and Trudgill 2004: 5).

This shows us that the mutual intelligibility seems to depend on geographical distances between the villages. But the question is does this traveller notice everything or just some changes. As it can be seen, a dialect continuum can be a range of dialects spoken in some geographical area that are only slightly different between neighbouring areas. But these differences become larger as someone travels in any direction, so that speakers from opposite ends of the continuum, are no longer mutually intelligible, they cannot understand each other. "A traveller walking in a straight line from village to village notices successive small changes, but seldom, if ever, observes large differences. This sounds like a justification of the continuum view, but there is an added twist. Might the traveller be misled by the perspective of most recent memory"(Heeringa and Nerbonne 2002: 375)?

For example, we can take numbers as representatives of dialects. Number one and two are closer to each other than one and three, and between one and ten there is no direct link whatsoever. But number five is in the middle, and it is as close to one as it is to ten. Do we say, in that case that one and ten are close? Or at least indirectly so? We can also compare the linguistic situation to colours. Pink and red can be two different colours, or pink can be thought as being a shade of red. But in that case is "red" also a shade of red, since "red" includes pink and what we think of as red. So, from that perspective there is no such thing as red, only shades of red. But on the other hand there is such a thing as red, since it is definitely not green or black. These analogies point to the fact that despite the difficulty, we do need some classification of linguistic data. We need to group them as closely as possible into dialects, keeping in mind that dialects interact and often shade off into each other forming dialect continua. There are in fact many such continua in Europe, such as the Romance, the Germanic, and the Slavic geographical dialect continuum. This of course entails a number of difficulties, such as how much difference is necessary to establish that two varieties are

two separate dialects rather than one single dialect, but we will turn to this issue below. As for dialect continua, in the next few sections we will explain some of the ways of how the concept of dialect and dialect continuum can be quantified and which instruments can be used to that end.

4.1. Cumulative Distances

Crucial for the operationalisation of the concept of dialect continuum are the idea of the cumulative, the Levenshtein, geographic, and phonological distances. These concepts are all intended to explain in more objective terms the idea of a dialect continuum. The cumulative distances are intended to explain cumulative distances/differences between dialects. We might explain it as follows: if we take three points X, Y and Z, which lie on a straight line, than we know that the distance from X to Z = the distance from X to Y + the distance from Y to Z. These distances can be calculated indirectly and directly. "If calculating indirectly, we can measure the distance via the intermediate points: $d(xn, x1)=d(xn, xn21)+d(xn21, x1)$" (Heeringa and Nerbonne 2002: 390-391). If we calculate directly, "we take the direct distance as it is given: $d(xn, x1)$"(Heeringa and Nerbonne 2002: 391).

This distinction can also be applied to phonological distances. If the phonological distances are calculated indirectly this means that the "traveller" remembers only the last variety so that he cannot compare the current variety to the varieties much earlier on the path.

4.2. Levenshtein Distance

There is another tool that is used to study and explore the dialect continuum, dialectometric method called: the Levenshtein distance. Chambers and Trudgill said that it is essential to the continuum view that these differences between certain areas are 'cumulative'. This means that the further the traveller goes, the larger the differences become, so that the mutual intelligibility depends on the geographic distance between certain dialects. This implies that the linguistic distance can be calculated very precise on the basis of geographic distance. If dialect were perfectly divided into areas then the traveller wouldn't notice any difference while he is within an area, and when he would enter another area he would notice a very big difference. But the dialects are not perfectly

divided into areas. Those are the reasons for studying the term ‘dialect area’, and the tool used for this study is the, already mentioned, Levenshtein distance. “The Levenshtein distance may be understood as the cost of (the least costly set of) operations mapping one string to another. The basic costs are those of (singlephone) insertions, deletions, and substitutions” (Heeringa and Nerbonne 2002: 380). Levenshtein distance gives us a more exactly approach to the issue of cumulatively, and helps us to get a measure of difference.

A simple, but good example is the pronunciation of “saw a girl” in Standard American and Boston.

> In Standard American “saw a girl” is pronounced as [sɔ:əgʌrl]. In Boston it is pronounced as [sɔ:rəg:l]. Now we can change the first pronunciation into the second as follows:
> sɔəgirl delete r 1
> sɔəgil replace i/ø 2
> sɔəgøl insert r 1
> sɔrtəgøl
> Total 4 (Heeringa and Nerbonne 2002: 381)

This is a very simple example that shows the fundamental idea. The simplest versions of Levenshtein distance are based on calculations of phonological distance. “Nonidentical phones contribute to phonological distance, whereas identical ones do not. Thus the pair [a,p] counts as different to the same degree as [b,p]. In more sensitive versions phones are compared on the basis of their feature values, so that the pair [a,p] counts as much more different than [b,p]” (Heeringa and Nerbonne 2002: 382). After reading and exploring the Levenshtein distance and its depending on phonological distance I conclude that we can apply this method on the whole language because I found many other examples (like the one “saw a girl”) where after calculating the phonological distance we got a measurement of similarity between two strings. For example “afternoon” is pronounced in one way in the dialect of Savannah, Georgia and in another way in the dialect of Lancaster, Pennsylvania. At the end, after deleting, inserting and replacing certain things we got a total 3. When we compare pronunciations in this way, we can conclude that the distance between longer pronunciations is generally greater than the distance between shorter pronunciations.

Another thing that I find “problematic” and it is connected with the Levenshtein distance is that the correlations between the Levenshtein distance distances and the perceptual are not perfect. It depends on the attitude of a listener towards the different dialects and also on their knowledge about the geographical position of the dialects. It can be seen that the Levenshtein distance is calculated only on the basis of lexical, phonetic and morphological material, not on syntax. This is another factor that makes the correlation imperfect.

4.3. Geographical Distance versus Phonological Distance

Geographical and phonological distances are related to each other. As cited in Chambers and Trudgill (2004: 5): "If we travel from village to village, in a particular direction, we notice linguistic differences which distinguish one village from another." This traveller develops a notion of indirect phonological distance (Figure 1). This distance is a sum of distances from neighbouring points on a connected line. We can see that the view of the traveller is misleading because the phonological paths do not sum along the geographical paths. The phonological distances are not cumulative, but the differences "accumulate in a linear fashion, giving the traveller the impression that the continuum is simple and dialectologically real" (Heeringa and Nerbonne 2002: 392). The traveller hears small differences, so that these pronunciation differences are a linear function of geography, but the true pronunciation difference simply is not the sum of the pair wise differences along the path.

Another thing that can be concluded is that lexical variables are not used when calculating the distances in dialect continuum. This leads to the conclusion that lexical variables are inherently difficult to define, as historically there has always been significant overlap in the vocabulary of certain varieties, and the traveller hears only the small differences in pronunciation, only certain changed phonemes within the 'same' lexeme.

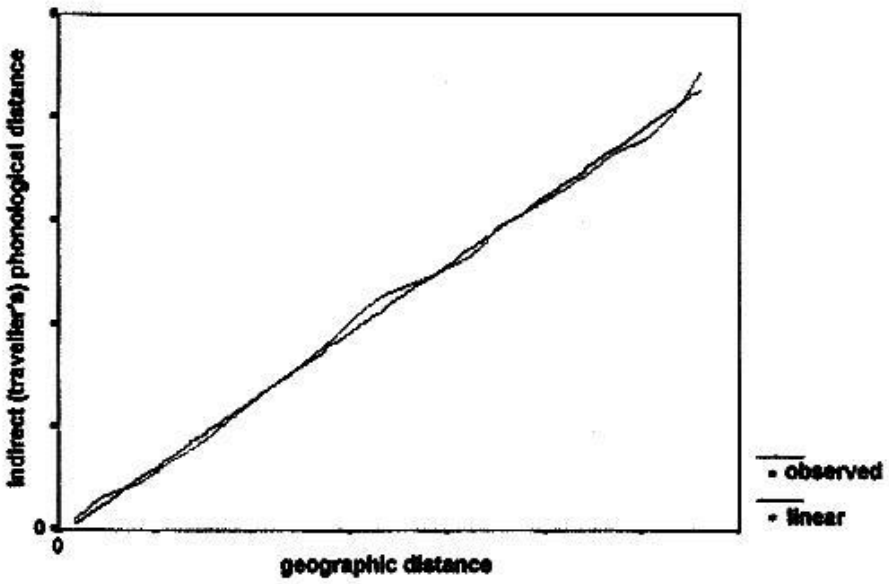

Figure 1. "Geographic distance versus mean indirect (traveller's) phonological distance. This graph explains the perception of Chambers and Trudgill's traveller that the dialect landscape is a simple accumulation of differences" (Heeringa and Nerbonne 2002: 393).

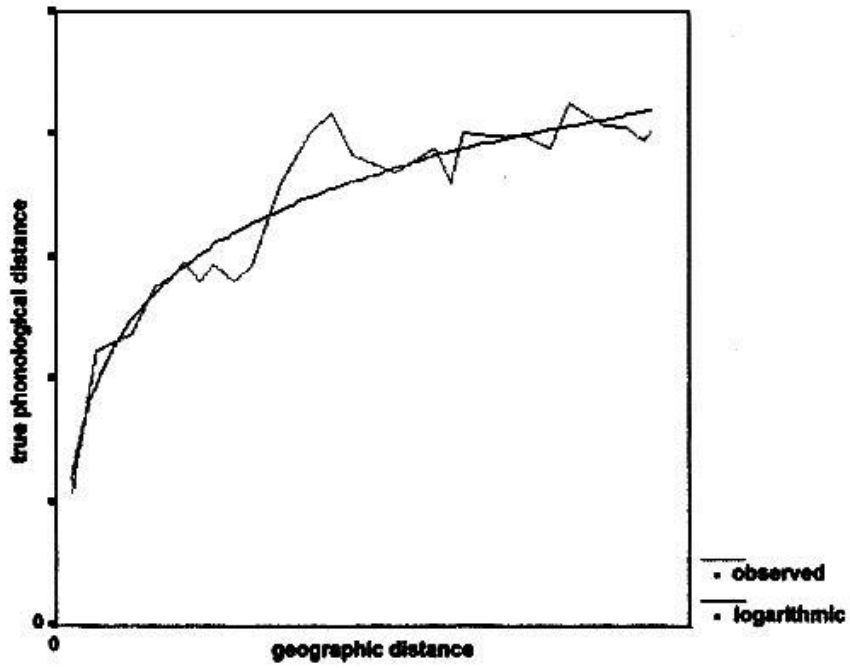

Figure 2. "Geographic distance versus mean true phonological distance. This graph illustrates the fallacy in the memory less traveller's view of the dialect landscape" (Heeringa and Nerbonne 2002: 394).

4.4. Isoglosses

Another term that must be mentioned while writing about dialect continuum is isogloss. Isoglosses are "the lines marking the boundaries between two regions which differ with respect to some linguistic feature (for instance, a lexical item, or the pronunciation of a particular word)" (Chambers and Trudgill 2004: 89). What this means is that this line is between two areas which disagree with each other, but share some linguistic aspects. A distinction between isoglosses and heteroglosses should me made. Isogloss is a single line that separates a region where feature X is found from the region where its counterpart Y is found. Heterogloss is formed by two lines that separate regions where X and Y are found. The lines link speakers with feature X and those with feature Y.

As already mentioned, dialectology has aimed to divide language areas into dialect areas mostly by drawing sharp borders (isoglosses) between the areas on a map. However, the use of isoglosses can cause a number of problems. Isoglosses are not always "perfect", they do not always coincide. They can run parallel, or even cross each other. The use of isoglosses also puts dialects in certain categories where no degrees of differences can be expressed. At the end, dialects might be dispersed by migration or war, and although they are related to each other they are not adjacent to each other anymore.

A very interesting bundle of isoglosses (the coincidence of a set of isoglosses) is the bundle that is dividing France into northern and southern region. This bundle is very significant because "not only does the bundle describe the location of important linguistic differences, but it also describes the location of venerable social and cultural differences" (Chambers and Trudgill 2004: 101-102). The bundle is at the same place where the 'southern' territory begins, and because of that the citizens to the south of the line consider themselves southerners and the citizens to the north, northerners. This feeling of allegiance is so strong because of the "ancient ethnic split between (as A. Brun put it) 'partially Romanised Celts in the north and thoroughly Romanised non-Celts in the south' (quoted by Jochnowitz 1973: 156)" (Chambers and Trudgill 2004: 102). We can also see some correspondence between the isoglosses bundle and non-linguistic boundaries. The bundle describes the division between biennial and triennial rotation of crops, an old split in legal practice that existed in France at the time of the fall of the Roman Empire. There was no actual line but it was in the same general area as the isogloss bundle, and in architecture the roofs of houses are differently built in the south and differently in the north, and the reasons are not functional, but

stylistic. This all shows how people are used to impress their individuality upon others, and that language contributes to this sense of community.

There is also a bundle of isoglosses in Massachusetts that divides it into eastern and western dialect, and that helped in the distribution of certain place- names. The eastern dialect is again with a bundle divided into north-eastern and south-eastern. The western region preferred place-names ending in –'field(s)', the north-eastern region also, although there are five uses of names ending in –'bury', and in the southeast the ending -'ham' is preferred. It can be seen how isoglosses, in a way, helped to determine the choice of place-names.

5. Geographical Dialect Continuum in Europe

To better understand dialect continuum it is crucial to name some of the most important ones. First we have the Western Romance dialect continuum.

> The standard varieties of French, Italian, Catalan, Spanish and Portuguese are not really mutually intelligible, but the dialects of these languages are all part of the Western Romance dialect continuum "which stretches from the coast of Portugal to the centre of Belgium (with speakers immediately on either side of the Portuguese–Spanish border, for instance, having no problems in understanding each other) and from there to the south of Italy (Chambers and Trudgill 2004: 6).

Another dialect continuum is the West Germanic one that "includes all dialects of what are normally referred to as German, Dutch and Flemish" (Chambers and Trudgill 2004: 6). Although this is one of the standard examples adduced in support of the concept of dialect continuum, there are grounds to believe that mutual intelligibility no longer is the case on the Dutch-German border. Once there was a dialect continuum, and mutual intelligibility of dialect speakers on both sides of the border was possible to quite some degree, but dialects on both sides of the border were too long under the influence of the standard languages of their respective nations, therefore it can be said that nowadays the dialect continuum is 'broken'. It is proven that today they communicate with each other more in English than Dutch or German.

The third one is the North Slavic Dialect continuum "including Russian, Ukrainian, Polish, Czech and Slovak, and the South Slavic continuum including Slovenian, Serbian, Croatian, Macedonian and Bulgarian" (Chambers and Trudgill 2004: 6).

5.1. Socio- political Implication of the Existence of a Dialect Continuum

It is very important to show how certain political and cultural factors created boundaries between different varieties of a dialect. "There is a great deal in literature on nationalism on the nation-language relationship" (Hoffman 1996: 32). A shared language is seen as something very important in the shared culture and economy (the modern French state, which has imposed a single language in order to unify the population) and a distinct language is important in demarcation one state of another (case in Central and East Europe in the 19th century, where limits of nations were linguistically defined). The demarcation of one language from another is complex, because we can find how related dialects or dialect continua have been divided into different languages without any linguistic reference. Here the important factors are the political and cultural ones. An excellent example is the Dutch-German dialect continuum already mentioned above in section 5.

> Many other language divisions between languages actually derive from long-standing political or religious-cultural divisions, rather than obviously linguistic cleavage; this is the case in the separation of Czech, Slovak and Polish, where some informants still claim a high degree of mutual comprehensibility, particularly between Czech and Slovak(Hoffmann 1996: 33).

Another good example is the Serbo-Croatian group, also mentioned in section 5, which belongs to the South Slavic continuum. Here we have several major dialects and three national standard languages: Bosnian, Croatian and Serbian. People are not used to thinking in continuum-type terms; they rather want to be absolutely separate. "Until recently, for example, Serbian and Croatian were thought of in Yugoslavia as a single language" (Chambers and Trudgill 2004: 7). Now they are separate countries, politicians want to stress their separateness, and the government of Bosnia wants Bosnia to have a third distinct language from the other two. "Similarly, Bulgarian politicians often argue that Macedonian is simply a dialect of Bulgarian – which is really a way of saying, of course, that they feel Macedonia ought to be part of Bulgaria" (Chambers and Trudgill 2004:7).

6. Social Dialect Continuum

After writing about geographical dialect continuum it must be said that dialect continuum can also be social and a good example for this is Jamaica. Jamaica has a very complex history. Briefly, "at one time the situation was such that those at the top of the social scale, the British, spoke English, while those at the bottom of the social scale, the African slaves, spoke Jamaican Creole" (Chambers and Trudgill 2004: 7). There is an extract form a poem in Sranan ("another English-based Creole spoken in Surinam") that shows that it is a language related to English but not mutually intelligible with it:

mi go – m'e kon	I've gone – I come,
sootwatra bradi	the sea is wide.
tak wan mofo	Say the words,
ala mi mati	you all my friends,
tak wan mofo	say the words.
m'go	I've gone,
m'e kon . . .	I come . . . (Chambers and Trudgill 2004: 7-8)

But over the centuries English influenced Jamaican Creole a lot. Jamaican Creole was considered an inferior form of it. At the end we have a result that "people at the top of the social scale speak something which is clearly English, and those at the bottom speak something which clearly is not, those in between speak something in between" (Chambers and Trudgill 2004: 8). This range from English to Creole is the social dialect continuum. There is also an example in West Indian dialect continuum where we can see that social dialect continuum depends on stylistic context:

"It's my book I didn't get any Do you want to cut it?

its mɑi buk ɑi didnt .et eni du ju wɔnt tu k_t it

iz mɑi buk ɑi didn .et non du ju wɑ_n tu kot it

iz mi buk ɑ din .et non ju wɑ_n kot it

ɑ mi buk dɑt ɑ in .et non iz kot ju wɑ_n kot it

ɑ fi mi buk dɑt mi nɑ bin .et non ɑ kot ju wɑ_n fu kot it" (Chambers and Trudgill 2004: 8).

We can at no place say where English stops and Jamaican Creole begins. "Again this is a difficult notion for many people to grasp, since we are used to thinking of languages as being well-defined and clearly separated entities: either it is English or it is not" (Chambers and Trudgill 2004: 9).

7. An Example of Dialect Continuum Shown on a Sample of 27 Dutch Towns and Villages

The best way to show how a dialect continuum exactly looks like is to show a sample that illustrates 27 Dutch dialects lying on a straight line. With Levenshtein distances the linguistic distance between all these dialects is calculated, then the relation between phonological and geographic distances is being researched, and it is also showed how dialect areas can be identified (this leads back to the idea of dialect continuum). "Finally, the multidimensional scaling is used to show how the dialects are related to each other" (Heeringa and Nerbonne 2002: 378).

Figure 3. "The locations of the 27 Dutch dialects studied" (Heeringa and Nerbonne 2002: 387).

7.1. Data of the Dialects

27 dialects from Dutch language area are being chosen. In the RND (*Reeks nederlands(ch)e dialectatlassen,*which was compiled by Blancquaert and Peé) "the same 141 sentences were recorded and transcribed for each dialect" (Heeringa and Nerbonne 2002: 378), and from these sentences 125 words have been chosen, and those words should represent "the range of sounds in the varieties" (Heeringa and Nerbonne 2002: 378)

Note: NOW = number of words for which more than one variant was used. Prof. =professions of the informants. Here we distinguished the following categories: a=agricultural, n=non-agricultural, s=student? = unknown. For housewives the profession of their husband was given. Ages = ages as given in the RND. For Scheemda, Veendam, Eext, and Beilen the birth dates were given. We calculated the ages by calculating the difference between the date of birth and the mean of the first year and the last year of the recording period. Period=recording period. Volume=part of the RND in which the dialect is found.

TABLE 1. *Information concerning the 27 dialects from the RND*

Place	NOW	Sex	Prof.	Age	Period	Volume
Scheemda	9	m/m	n/n	64/72	1956–1961	16
Veendam	4	m/m	n/n	69/62	1956–1961	16
Eext	7	m/m/f	a/a/a	57/58/61	1956–1961	16
Driebergen	2	f/m/m/m	n/n/n/n	50/81/80/59	1950–1962	11
Koekange	0	?/m	?/a	76/73	1974–1975	14
Hasselt	0	m/m	n/a	62/66	1974–1975	14
Staphorst	1	m/m	a/a	69/47	1974–1975	14
Zalk	0	f/m	?/a	52/56	1974–1975	14
Oldebroek	0	f/m	?/n	53/32	1974–1975	14
Nunspeet	1	m/m	n/a	50/53	1950–1970	12
Putten	7	f/f/f	n/a/a	38/28/54	1950–1970	12
Amersfoort	4	m/f	n/n	71/58	1950–1970	12
Beilen	5	f/m	a/n	74/36	1956–1961	16
Ruinen	0	f/f/f	?/?/?	59/67/65	1974–1975	14
Ossendrecht	2	m/f/m	n/n/n	63/22/18	1933–1935	3
Clinge	0	m/m/m	n/s/s	39/13/12	1933–1935	3
Moerbeke	0	m/m/m	s/n/n	23/20/54	1933–1935	3
Lochristi	1	m/m/m	n/n/n	52/29/48	1933–1935	3
Vianen	1	m/m/m	a/n/a	66/30/61	1950–1962	11
Hardinxveld	1	m/m	n/n	77/80	1939–1949	9
Zevenbergen	0	m/m/m	n/?/n	36/79/41	1939–1949	9
Oudenbosch	1	m	n	46	1939–1949	9
Roosendaal	0	m/f/m	n/n/n	63/63/27	1939–1949	9
Bellegem	4	m/m	n/n	35/69	1934–1940	6
Nazareth	0	m/f/m	n/n/s	25/20/24	1927–1930	2
Waregem	4	m/m	n/n	77/63	1934–1940	6
Zwevegem	0	m/m	n/n	35/33	1934–1940	6

It is almost impossible to draw generalizations about the dialect gradation from such information, "however, this is possible when using the Levenshtein distance measured over large samples of vocabulary" (Herringa and Nerbonne 2002: 380). If we assume that 125 words are transcribed from two different dialects then the Levenshtein distance can then be calculated for 125 word pairs. "The total distance between the dialects is equal to the sum of the 125 Levenshtein distances" (Heeringa and Nerbonne 2002: 383).

8. Dialect Continuum of Ancient Greek

Another interesting, and very good example of dialect continuum is the dialect continuum of Ancient Greek. Some basic principles of analysis of dialectal picture will appear in the case of Ancient Greece and this will lead to the influence of synchronic and diachronic approach on the classification and interpretation of certain data and information. Examples that will illustrate some other basic principles, introduced earlier in the text, will also appear while talking about dialect continuum of Ancient Greek.

When we talk about dialects of Ancient Greek and about Greek dialectology in general, the problem is always the Greek classification. "The whole history of the study of the Greek dialects bears witness to the fact that a clear-cut classification can only be achieved by emphasizing one set of linguistic features in a given dialect at the expense of others" (Finkelberg 1994: 2). There are no clear-cut demarcation lines between dialects, but to make a classification in order to classify certain dialects there are two approaches mentioned. The first is characteristic of the comparative method where we have certain "mixed" or "impure" models, which do not correspond to the "pure" models. The second approach that has been adopted in dialect geography says that these "pure" models of dialects are nothing but a convention, and that we are talking about gradual changes which form dialect continuum. This approach was developed from the studies of living languages, like French, German, and Italian. Barely any boundaries can be drawn, but "a really clear-cut idiom, such as that presupposed by the comparative method, can only emerge as a result of a split of the dialect continuum caused by external factors such as migration or foreign invasion" (Finkelberg 1994: 2).

When talking about ancient Greek two migration processes that need to be mentioned are Dorian invasion of the Peloponnese and migrations to the East, to Asia Minor. But it still can be

seen that these dialects are actually fragments of a dialect continuum which existed before the period of these great migrations. "Dialect geography of the languages that have developed under settled conditions naturally proceeds from the map, and the charts drawn in this kind of dialect geography are normally the result of empirical observation" (Finkelberg 1994: 5). This is the reason why it is said that dialect geography is often misleading and has only limited applications.

8.1. Synchronical Approach

In this section and the next section, 8.2., the goal is to show that orientation of the analyst itself - synchronic or diachronic, can affect the contour of a dialectal image. It can be said that certain dialects are forming a continuum when the relationship between those dialects (A, B, C, D, E) are of the type ab, bc, cd de. This means that there is a continuum when we have combinations of some linguistic features which can also be found in other dialects, and not a unique idiom. "The only units which cannot be reduced to the other ones would be those found at the opposite extremes of the continuum (a in ab and e in de)" (Finkelberg 1994: 5). If we take the dialect B, it proves to be reducible to dialect A and dialect B at the same time. The same thing can be said for dialect C, being between B and D, and for dialect D being between C and E. Eventually there is the end of the continuum, and this unit cannot be reduced to two others.

This will finally be shown on concrete Greek examples. Finkelberg put 6 Greek dialects, according to the principle above, (Attic)-Ionic, Arcado-Cyprian, Lesbian, Thessalian, Boeotian, and West Greek as a whole. Lesbian was shown to have features in common with both Ionic and Thessalian, so they decided to begin the test by placing Lesbian between those two dialects.

As I have already mentioned Lesbian has linguistic features in common with both Ionic and Thessalain. Some of those features are oí/αi as nom. pl. of the article, τ in temporal conjunctions of the ὅτε type, and the Lesbian infinitive ending of athematic verbs relates to both, the Ionic and the Thessalian .It can be seen that when Lesbian is placed between Ionic and Thessalain, those three dialects form a dialect continuum.

There are linguistic features where Thessalian is isolated from Lesbian and Ionic, and these features point to the direction of Boeotian. Some of the features that the three dialects have in common are o for α with liquids, πε- from *kŵe-, the same suffixes in dat. pl. of consonant stems,

and the conjunction αί and the particle κε/κα. To conclude, when Thessalian is places between Lesbian and Boeotian we can say that they form a dialect continuum.

Linguistic features of Boeotian that cannot be found neither in Thessalain or Lesbian, point in the direction of West Greek, so that the dialect Boeotian can be put between Thessalain on the one side and West Greek on the other. Some of the features that these three dialects have in common are the vocalism e in the verb "to wish, the conditional conjunction αι, the modal particle κε/κα, έν with the accusative (North-West Greek only), etc

Arcado-Cyprian was the dialect that was mentioned at the beginning but wasn't placed anywhere between those dialects in a dialect continuum. Arcado-Cyprian is one of the dialects in which direction Ionic points. Now, this means that Arcado-Cyprian stands before Ionic, first in the scale or that it should be placed in another place in the continuum. To put Arcado-Cyprian in first position or after Ionic proves wrong, but to put it on the place of Lesbian, which means between Ionic and Thessalian, proves to be the right option. At the end we have a continuum that looks like this: West Greek ←→ Boeotian ←→ Thessalian ←(Lesbian → Ionic

A.-Cypr.)

8.2. Diachronical Approach

The diachronical point of view gives a historical development and changes of those dialects. Three dialect areas can be isolated according to three distinct groups of linguistic features:

(a) Only Ionic, Arcado-Cyprian and Lesbian share some same features so that it can be concluded that dialect within this isoglosses (Ionic, Arcado-Cyprian and Lesbian) can form a single dialect area.

(b) Arcado-Cyprian, Lesbian and Thessalian share the same isoglosses, ν for ο, and the athematic inflection of contract verbs. Here can be concluded that that dialect within this isoglosses in question (Arcado-Cyprian, Lesbian and Thessalian) form a single dialect area.

(c) Lesbian, Thessalian and Boeotian share the also the same isoglosses, dat. pl. of consonant stems, and some other features. It can be concluded that dialects within the isoglosses in question (Lesbian, Thessalian and Boeotian) form a single dialect area.

It can be seen that Lesbian is within each of the three groups of the isoglosses, and Arcado-Cyprian and Thessalian in two, which shows that there is some overlapping between the dialect areas, and the dialect would form a continuum even if not being grouped into dialect areas.

It can also be said that those three dialect areas relate to each other chronologically, and we can describe this in three stages:

> Stage One – differentiation off the dialect area comprising the future Ionic, Arcado-Cyprian and Lesbian and opposed to the dialect area which comprised the future West Greek and Boeotian; Thessalian in contact with both
>
> Stage Two- differentiation of the dialect area comprising the future Arcado-Cyprian, Lesbian and Thessalian and opposed both to West Greek (Boeotian in contact with both) and, significantly, to Ionic, isolated at this stage as an independent dialect
>
> Stage Three - differentiation of the dialect area comprising the future Lesbian, Thessalian and Boeotian and opposed to West Greek, to Ionic and, significantly, to Arcado-Cyprian, isolated at this stage as an independent dialect (Finkelberg 1994: 22-23).

The positions of those Greek dialect can be evaluated only through two approaches: synchronical and diachronical, and the analysis showed at the end "that it is likely that Stage Two should be placed between 1600 and 1200 B.C. and Stage Three between 1200 and 900 B.C" (Finkelberg 1994: 31). If Stage One was put between 1900 and 1600 B.C., this would mean that this particular dialect continuum of Ancient Greek began with the coming of the Greeks into Greece in the Middle Bronze Age and ended with the great migrations in the Early Iron Age.

9. Conclusion

All thing considered, it can be seen that it has always been difficult to decide what a language and what a dialect is, and that autonomy and heteronomy of a dialect or language are more a result of cultural and political than linguistic factors.

When a range of dialects are spoken in some geographical area that are only slightly different between neighbouring areas we talk about dialect continuum, and dialect continuum cannot be discussed without mentioning the cumulative Levnshtein, geographic, and phonological distances which help in measuring the differences between certain dialects. It can also be concluded that the boundaries (isoglosses) between certain regions, which differ with respect to some linguistic feature, can cause a number of problems because isoglosses are not always "perfect", they do not always coincide, and the use of them also puts dialects in certain categories where no degrees of differences can be expressed. Certain political and cultural factors created boundaries between different varieties of dialect.

All this has been illustrated on 27 Dutch dialects lying on a straight line where with the Levensthein distances the linguistic distance between all these dialects is calculated. Another illustration was the dialect continuum of Ancient Greek where some basic principles of analysis of dialectal picture appeared. 6 dialects of Ancient Greek form a continuum that looks like this: West Greek ←→ Boeotian ←→ Thessalian ←(Lesbian → Ionic, and on this example it has been

A.-Cypr.)

proven that a continuum appears when we have combinations of some linguistic features which can also be found in other dialects, and not a unique idiom.

Some basic problems and principles in the interpretation of the dialectal data have been explained, and it has been shown how actually the whole language depends on social, political, historical and cultural changes.

10. References

Chambers, J.K., and Peter Trudgill (2004). *Dialectology.* 2nd edition. Cambrigde: Cambridge University Press.

Clachar, Arlene (2003). Paratactic conjunctions in creole speakers' and ESL learners' academic writing vol.2: 271-289. Blackwell Publishing Lid.

Finkleberg, Margarit (1994). *The Dialect Continuum of Ancient Greek* vol.96: 1-36. Department of the Classics, Harvard University is collaborating with JSTOR.

Heeringa, Wilbert, and John Nerbonne (2002). *Language Variation and Change: Dialect areas and dialect continua* 375-400. Cambridge: Cambridge University Press.

Hoffman, Charlotte and the authors of individual chapters (1996). *Language, Culture, and Communication in Temporary Europe:* 28-61. British Library Catalogue in Publication Data.

Romaine, Suzanne (2000). *Language in Society: An Introduction to Sociolinguistics.* 2nd edition. New York: Oxgord University Press.