The Origin of the Article in Indo-European Languages of Western Europe

Elena Andreevna Makarova

Senior Lecturer, Educational Center "Logos Express", 31, Pokrovka Street, Moscow, Russia

Vladimir Nikolaevich Polyakov

*Candidate of Technical Sciences, Senior Researcher of the Institute of Linguistics of Russian Academy of Science
1/12, Bolshoy Kislovsky lane, Moscow, Russia; pvn-65@mail.ru*

Abstract

This paper is concerned with the origin of the article in Indo-European languages of Western Europe. Several hypotheses concerning the origin of the article are studied, including the hypothesis of spontaneous and independent development, the Arabic-origin hypothesis, the Hebrew-origin hypothesis and the Biblical-origin hypothesis. We suggest that the main source of article borrowing into the ancient languages of Western Europe (Germanic and Romance) was the Bible. Supposedly, the grammatical category in question penetrated into the languages when the Bible was translated into national languages. We present a historical analysis of literary monuments in Old French, Old Spanish, Old German, and Old English. This shows that these languages had acquired the article before the Bible was translated into the mentioned national languages. It allows us to suppose that Ulfilas' Gothic Bible, which appeared earlier, was the source of penetration of the article into Western European languages. This assumption is based on the analysis of literary monuments in ancient languages spoken in Europe, as well as on the comparison of the geographical spread of the article in European languages and the map of Gothic conquests in the 6th century AD. Some of the research draws upon the electronic linguistic resources WALS (http://wals.info), the "Languages of the World" database of Institute of Linguistics of Russian Academy of Sciences, and the data of ASJP project (http://asjp.clld. org/).

Keywords: Article, Gothic, Indo-European, Bible, Ulfilas

1. Introduction

The article as a grammatical category of a language is tightly connected with its case system (system of actant relations), fixed word order, and topic-comment information structure of the sentence. The article is a widely used category in modern Indo-European languages of Western Europe (English, German, Spanish, and Portuguese). In contrast, this category is almost completely absent from all Slavic Indo-European languages (except Macedonian and Bulgarian) and from Uralic and Mongolian languages (Fig. 1). The development and spread of the article in Indo-European languages is intriguing, since Proto-Indo-European languages lacked it. Thus, the article does not exist in ancient Indo-European languages such as Old Persian, Avestan, and Latin. We will discuss several hypotheses of the origin of the article, including the traditional hypothesis of spontaneous and independent development (Greenberg, 1978), the Arabic-origin hypothesis, the Hebrew-origin hypothesis and the Biblical-origin hypothesis. We suggest that the main source of article borrowing into the ancient languages of Western Europe (Germanic and Roman) was the Gothic Bible translation.

The translation of the Bible into national languages could have given impetus to the penetration of the article into western Indo-European languages—Old French, Old Spanish, Old German and Old English in particular. However, our historical analysis of literary monuments in the mentioned languages shows that these languages had acquired the article before the Bible was translated into them. It allowed us to suppose that it was specifically Ulfilas' Gothic Bible (written in the middle of the IV-th century) which made Gothic the first language to borrow the article from Koine Greek, and that Gothic became a mediator in the process of borrowing of the article from Koine Greek into Indo-European languages spoken in Europe during the Middle Age (along with the Vulgate, Ulfila's Bible was one of the first translations of the Bible into national languages – Latin and Gothic).

This assumption is based on the analysis of literary monuments in ancient languages spoken in Europe, as well as on the comparison of the geographical spread of the article in European languages and the map of Gothic conquests in the 6th century AD.
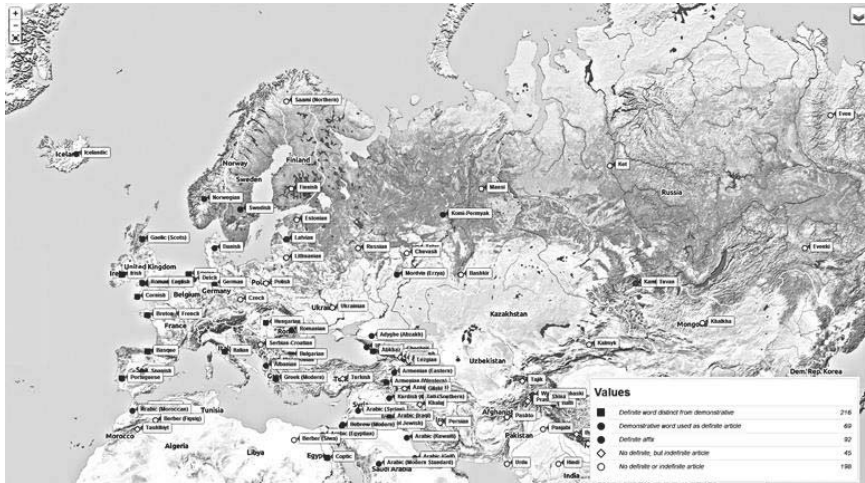
Fig. 1 Map of the distribution of the definite article in Eurasian languages (Dryer, 2013)

Following the introduction to this paper is a section dedicated to the function of the article in general, an excursus into the history of the development of the article in Indo-European languages, an overview of the hypotheses on how the article could be borrowed into Indo-European languages of Western Europe, and, finally, a section with conclusions.

2. The Main Function of the Articles and Other Means of Its Realization

The article is a part of speech, the main function of which is to express the definiteness of the word it refers to (be it a noun, a substantiated adjective, a nominalized verb, a nominalized numeral or a nominalized participle). The articles are divided into several groups, however, not all of these groups must necessarily be present in one language:
- Definite article (indicating that the noun, or, to be more precise, a concept, denoted by a noun, is identifiable to the listener);
- Indefinite article (indicating an unknown object or person, one representative of a group of similar objects);
- Zero article (occurring in noun phrases that contain no article);
- Partitive article (denoting a part of something uncountable).

Some languages, such as French, German or Italian, also have a contracted form produced by combinations of certain prepositions and an article. For example, in French: the preposition *à* and the definite masculine article *le* become *au*. Or in German: the preposition *in* and the definite article *das* become *ins*.

Articles may be sensitive to the same grammatical categories of the nouns they modify. In English, which has lost the grammatical categories of gender and case, there is only one form of the indefinite article '*a*' and one form of the definite article '*the*'. In French and Spanish, where nouns have two genders (feminine and masculine), both the definite and the indefinite articles have a feminine and a masculine form. In German, which has three genders and four cases, definite and indefinite articles have a masculine, a feminine and a neuter genders, as well as case forms.

There is no argument about the derivation of the modern forms of the article in Western Indo-European languages. In these languages the indefinite article developed from the numeral 'one', either preserving the same form, e.g. *un* in French, *ein* in German, or changing into a different word, e.g. *a* in English. The definite article derives from demonstrative pronouns: in Romance languages (French, Spanish) they come from the Latin demonstratives *ille* and *illa* (Brachet, 1876). The English definite article *the* developed from the demonstrative pronoun *þe* in Middle English (Hoad, 1996), and in German the definite articles derived from Old High German *ther* (Bisle-Müller, 1991).

As mentioned above, the main function of the definite article is to express definiteness of the noun or the noun group it modifies. Definiteness is a feature of a noun or a noun phrase, which serves to distinguish between specific and identifiable entities and entities that are not identifiable.

Nevertheless, there are languages that do not have a grammaticalized concept of definiteness, i.e. an article. They may instead have a variety of other means of expressing definiteness. Let us look at them more closely.

1) Lexical means, such as:

- demonstrative pronouns (they specify the object that is being spoken about, the location of this object relatively the speaker or the addressee);
- possessive pronouns and nouns in the possessive case (they make the object or the person identifiable in the given context by denoting their owner);
- numerals (they refer to the nouns that have already been mentioned and are known to the listener).

2) Prosodic means:

Sometimes, intonation may be the only means of marking a definite object and an indefinite one. The unknown or the indefinite is always stressed, while the definite remains unstressed. For example, in Russian (ex.1).

(1) Мальчик пришел.
(Malchik prishol).

If the first word is stressed, it will be translated as 'A boy has arrived'; if the second word is stressed, the sentence will mean 'The boy has arrived'.

3) Topic-comment, or word order:

In languages that do not have a fixed word order, the topic, or something that is already known and definite for the listener, is placed at the beginning of the sentence, and the comment, or new information, at the end. Nevertheless, the sentence can begin with the comment, but in this case the noun (or the noun phrase) is accompanied by an indefinite pronoun or adjective, like 'some' (ex. 2, Ukrainian).

(2) Якийсь чоловік зайшов до кімнати.
(Jakis' cholovik zaishov do kimnaty)
'Some man entered the room'.

In this case the sentence begins with the comment, and it is preceded by an indefinite pronoun, which shows that the noun 'man' is indefinite, as it is mentioned for the first time.

4) Context:

In some cases there are no indicators of definiteness or indefiniteness of the noun in the sentence (ex. 3, English).

(3) A woman bought the book.

In languages that have no articles, both the nouns 'woman' and 'book' lack indicators of definiteness or indefiniteness, and only the context of the sentence can help identify which of them is definite, and which is not.

5) Cases:

The case of a noun can serve as a means of expressing definiteness. For example, in Russian the genitive case refers to something indefinite, and the accusative case – to something definite (ex. 4, English).

(4a) I didn't write a/the letter.

In Russian, which does not have the article, the difference between 'a letter' and 'the letter' is expressed by the case of the noun (ex. 4a).

(4b) Я не писал письмо/письма.
(Ja ne pisal pis'mo (accusative)/pis'ma (genitive)

3. History of the Development of the Article in Indo-European Languages

The article is one of the features characteristic of Indo-European languages, as it occurs in over 50% of Indo-European languages (see Fig. 2).

In order to construct the map in Fig. 2 we used a relatively new method of contrast queries, first described in (Anisimov, 2013 & Solovyev, 2013).

The query used the database "Languages of the World" of Institute of Linguistics of Russian Academy of Sciences, as well as the program LangFamilies, written in VBA, which calculates the frequency of use of grammatical features in the language families, branches and groups that are present in the database "Languages of the World" of IL RAS. The calculation results are stored in a MS Excel table. We used the interface of MS Excel to process queries on any

combination of frequencies and to form sets of grammatical features that meet the query. The search master of the database "Languages of the World" of IL RAS formed sets of the languages that have the given number of features. We used the following query for the table LangFamilies: "To find features that are present in minimum 50% of Indo-European languages and in maximum 5% of Altaic languages."

Each line in the query (the result of the query in the LangFamilies program is used as a query in the database "Languages of the World" IL RAS) shows one grammatical feature that is present in the contrast query to the stated subset of languages. As the database is organized in a hierarchal way, each line shows the grammatical feature at the end of the chain. Other features are shown for information on the branch of the tree the feature belongs to.

The number in brackets, e.g. (585), means the inventory number of the feature in the database. The vertical line (|) separates an inferior branch from the superior branch; the number of dots before the name of the feature means the level in the tree. The phrase "IS PRESENT" means that the feature it refers to is present in the query.

The features relevant to some language family are a set of features characteristic of over 50% of these languages and less than of 5% in another (contrast) family. For example, in contrast to Altaic, Indo-European languages have the following set of features (Table 1).

Table 1. Set of features relevant for Indo-European languages in contrast to Altaic

SIGN (584) 2.3.0.MORPHOLOGICAL TYPE OF LANGUAGE|(585) .METHOD OF JOINING MORPHEMES IN WORDS|(591) ..INFLECTIONAL LANGUAGES IS PRESENT
SIGN (639) 2.3.2.SUBSTANTIVE CLASSIFICATIONS|(641) .CLASSES OF AGREEMENT IS PRESENT
SIGN (639) 2.3.2.SUBSTANTIVE CLASSIFICATIONS|(641) .CLASSES OF AGREEMENT|(642) ..GENDER IS PRESENT
SIGN (639) 2.3.2.SUBSTANTIVE CLASSIFICATIONS|(641) .CLASSES OF AGREEMENT|(646) ..MEANS OF EXPRESSING CLASSES OF AGREEMENT IS PRESENT
SIGN (639) 2.3.2.SUBSTANTIVE CLASSIFICATIONS|(641) .CLASSES OF AGREEMENT|(646) ..MEANS OF EXPRESSING CLASSES OF AGREEMENT|(652) ...MORPHOLOGICAL IS PRESENT
SIGN (639) 2.3.2.SUBSTANTIVE CLASSIFICATIONS|(641) .CLASSES OF AGREEMENT|(646) ..MEANS OF EXPRESSING CLASSES OF AGREEMENT|(653) ...SYNTACTIC IS PRESENT
SIGN (639) 2.3.2.SUBSTANTIVE CLASSIFICATIONS|(641) .CLASSES OF AGREEMENT|(646) ..MEANS OF EXPRESSING CLASSES OF AGREEMENT|(653) ...SYNTACTIC|(661) ....IN ADJECTIVES IS PRESENT
SIGN (639) 2.3.2.SUBSTANTIVE CLASSIFICATIONS|(641) .CLASSES OF AGREEMENT|(663) ..MOTIVATION OF CLASSES OF AGREEMENT IS PRESENT
SIGN (639) 2.3.2.SUBSTANTIVE CLASSIFICATIONS|(641) .CLASSES OF AGREEMENT|(663) ..MOTIVATION OF CLASSES OF AGREEMENT|(672) ...GENDER IS PRESENT
SIGN (1290) 2.3.6.DEICTIC CATEGORIES|(1369) .ORIENTATION OF ACTION IN SPACE|(1370) ..EXPRESSION|(1378) ...PREPOSITIONS IS PRESENT
SIGN (1290) 2.3.6.DEICTIC CATEGORIES|(1380) .DEFINITENESS/INDEFINITENESS OF SUBSTANTIVE|(1382) ..ARTICLES IS PRESENT
SIGN (1439) 2.3.7.PARTS OF SPEECH|(1470) .SYNCATEGOREMATIC WORDS|(1477) ..PREPOSITION IS PRESENT
SIGN (1484) 2.4.0.PARADIGMS|(1619) .ADJECTIVE|(1620) ..INFLEXION|(1627) ...GENDER IS PRESENT
SIGN (1484) 2.4.0.PARADIGMS|(1638) .NOUN|(1661) ..CLASSIFYING CATEGORIES|(1664) ...GENDER IS PRESENT
SIGN (1735) 2.5.2.WORD DERIVATION|(1736) .DERIVATION|(1737) ..AFFIXATION|(1739) ...PREFIXES IS PRESENT

The location of Indo-European languages that satisfy the query in the database "Languages of the World" of Institute of Linguistics Russian Academy of Sciences is shown in Figure 2.
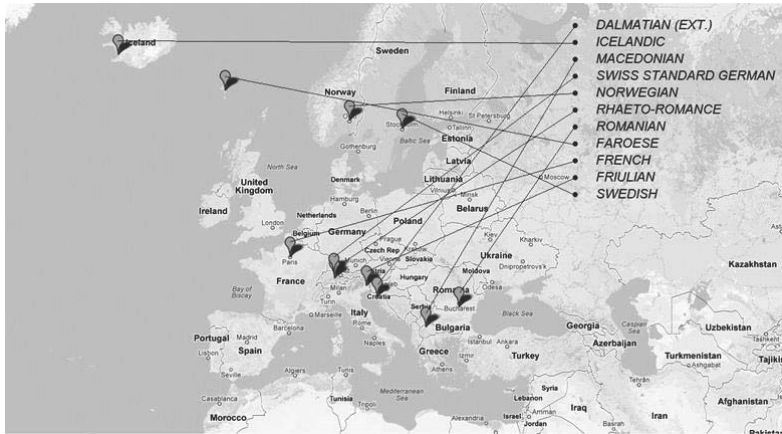
Figure 2. Map of Indo-European languages which have the set of relevant features (Google Maps is used for map visualization)

The set of relevant features and the set of languages that possess them were extracted from the database "Languages of the World" IL RAS by the method described above. At the same time they are relevant, but not genetic features. The map in Fig. 2 does not show all Indo-European languages, but only those that have the full set of the relevant features. As we can see, the map does not include extinct languages (Avestan, Old Persian, Latin, Polabian). Then we weaken the initial query by excluding some features: (1382) ARTICLES NOT PRESENT, (1739) PREFIXES NOT PRESENT. As a result, a larger set of languages emerges. List of languages added to the initial set:

- Istriot
- Provencal
- Spanish
- Avestan (ext.)
- Belarusian
- Gothic (ext.)
- Old Persian (ext.)
- Ishkashimi
- Latin (ext.)
- Polabian (ext.)
- Russian
- Sorbian
- Slovene
- Old Church Slavonic (ext.)

As we can see, the article was initially not part of the grammatical structure of old Indo-European languages (Latin, Old Persian, Avestan). Within the phylogeny of Indo-European, Slavic languages separated earlier than Western Indo-European ones (Germanic, Romance), and they preserved the feature of the lack of the article.
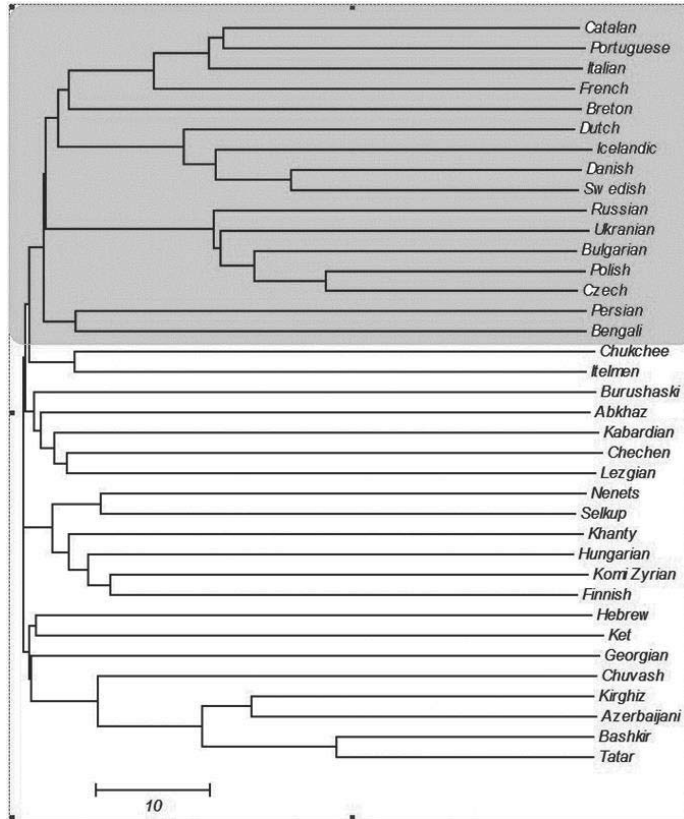
Figure 3. Tree of lexical similarities, including a sub-tree of Indo-European languages (highlighted), created through ASJP methodology (Polyakov et al., 2009). Note that this tree also shows some contact effects, such as the positioning of Breton.

This situation raises interesting questions, since it is highly improbable that the grammatical category of the article spontaneously appeared and developed in one part of kindred languages across Western Europe and, at the same time, that this category should have dropped out across Eastern Europe. It is also important that these changes took place in the period when written languages already existed, i.e., in historical time, and, thus, can be traced by literary monuments and correlated with historical events of early Christianity.

The combination of these facts, namely:

- lack of the article in the Proto-Indo-European language;
- rapid and massive penetration of the article in Indo-European languages of Western Europe in the period of early Christianity;
- non-penetration of the article in Indo-European languages of Eastern Europe;
- existence of historical chronicles and literary monuments of that time,
- allows us to suggest the hypothesis that the article did not penetrate into Indo-European languages of Western Europe accidentally, but rather under the influence of some particular linguistic situation, which we shall proceed to discuss.

4. Possible Hypotheses on the Way the Article Penetrated in Indo-European Languages

4.1 The spontaneity hypothesis

This is the most popular hypothesis on the development of the article in languages, introduced by Joseph Greenberg (Greenberg, 1978). As he claimed, the articles developed in several language families spontaneously and independently (Greenberg, 2004: 460).

From our point of view, this hypothesis is problematical on probabilistic grounds. The map (Fig. 4) shows the geographical spread of the article in the languages of Europe. If the process of creation and development had happened spontaneously and independently, the distribution in Europe would look differently, resembling something like a chess-board, where areas with the article would alternate with areas without the article. However, the article dominates in the West, but it is absent in the East. We believe that the simultaneous emergence of the article as a separate grammatical category from demonstrative pronouns in a number of languages from different branches of a language family must have been stimulated by a common linguistic situation: one linguistic center dominating the others could be the source of borrowing of the article.

Three variants of this hypothesis are possible: the Arabic-origin, the Hebrew-origin, and the Biblical-origin hypothesis.

The spontaneity hypothesis was applied to Indo-European languages in Bauer (2007), a study which we return to in the Discussion section.

4.2 The Arabic-origin hypothesis

This and the following hypothesis belong to the authors of the present study.

As is well known, Arabic has a definite article *al-* (Ryding, 2005). This could have penetrated into Indo-European languages during the Muslim Conquests.

The Arabic-origin hypothesis claims that during the Muslim conquest, which began in the 7th century, the Arabic language influenced the grammatical and lexical structure of the language spoken on the territory of modern Spain. And, later on, these changes penetrated into the dialects of the peoples who lived in Western Europe (modern French, German, Portuguese, English, etc.).

At the beginning of the 7th century, having conquered Arabia, Syria and Egypt, Arabs moved to the Iberian Peninsula. In 712 Iberia became an Islamic state under the government of the Umayyad Sovereignty. In 714 Arabs went to the western Basque region and, facing no opposition, soon reached Gallaecia. Nevertheless, the steep western and central Pyrenean valleys remained unconquered. During the same short period of time (714-716), the principal urban centers of the Iberian Peninsula surrendered (Lomax, 1978: 15-16).

The life of people did not change much during the two centuries following the beginning of the Muslim invasion (Collins, 1989: 39-50). In 713, the Visigothic count Theodomir signed a capitulation statement, according to which his lands became an authoritative state under Umayyad rule. In exchange of taxes, Arabs promised to respect the Gothic government and Christians. Lots of other Iberian towns followed this example, so the dwellers kept living under the Visigothic Law Code, being allowed to practice their faith. Thus, the Christian Church remained up until the end of the 8th century, and Latin remained the official language until the 11th century.

Historical linguistics provides plenty of evidence of the influence of Arabic on the modern Spanish language (Quintana, 2002) and (Lapesa, 1942).

Thus, according to the Arabic-origin theory, articles should first have appeared in the language of the people who inhabited the Iberian Peninsula in the 8th century and later, and after that it could be borrowed by peoples living further to the north-east.

Nevertheless, the earliest literary monuments of the German language, show that as early as in the second half of the 8th century the German language already had a system of articles, for instance, articles are found in "Mersebunger Zauberspruche" ("Mersenburg Incantations", circa 750) the earliest literary monument of the Old High German language (The complete list of all studied literary monument and examples from them is available online at https://cloud.mail.ru/public/Hni7/ZyUgvdi7p). This language was spoken by Germans living on the territory of the Frankish Kingdom, which was not conquered by the Muslims and, thus, could not be influenced by the Arabic language to any considerable extent.

In short, the Arabic-origin hypothesis is untenable.

4.3 The Hebrew-origin hypothesis

Biblical Hebrew has a definite article "ha" (Sáenz-Badillos, 1993). This could have been borrowed from Hebrew in two ways. Either people inhabiting the territory of Western Europe in the Early Middle Ages borrowed it by communicating with Jews who lived among them. And later on the article was fixed in the written from in different religious texts, poetry, etc. Or the article could have been borrowed from Old Hebrew following the translation of the Old Testament into national languages. In fact, this sub-hypothesis is a variation of the Biblical hypothesis (see below).

4.3.1 Jewish settlement

The article could have been borrowed from Old Hebrew when the Jewish people settled in Europe in the early Middle Ages. Old Hebrew, or Biblical Hebrew, had a definite article – prefix *ha*-. There is very little documentation regarding the life of Jews in the early Middle Ages, but we can speak about three centers of Jewish settlements in Christian Europe: Italy, the Balkan Peninsula, and Spain. These were the regions where the power of the Great Roman Empire was the most developed (Roth, 1994). After the fall of the Roman Empire in the 5th century, Europe was conquered by the Goths. At first they followed the Arian direction in Christianity, but after 586 they adopted Catholicism. This is when the life of Jewish people in Europe changed completely. The Visigoths wanted to establish an exemplary Christian society, thus, they tried to limit all the possible economic and social connections between Christians and Jews (Furst, 1849). The theological principle of the Christian religion did not permit converting Jews into Christianity by force. Jews were living witnesses and proof of the antiquity of the Holy Script. Another reason why the Christian Church tolerated Jews was that the humbling position of this despised minority constantly reminded them about the mistake they had made and of the fact that God had rejected them. It was believed that when the world ends, Jews would accept the truth of the Christian religion and adopt it, and it would be the final stage of the victory of Christianity.

Thus, in the early Middle Ages the Jews were not plenipotentiary citizens of Europe. They were not allowed to hold administrative positions and to have Christian slaves. Their basic attitude was that they were a minority group, living in Europe as in an asylum that the God condemned them to. They surrendered to their life conditions and made no attempts to change the current state, treating it as a temporary punishment for the past sins.

Thus, the assumption that Old Hebrew influenced the use of the article in Europe during linguistic contacts seems improbable. We would rather expect Hebrew to be a recipient than a donor of linguistic influence. Among the compact groups of Jews living in linguistically different foreign areas, different dialects of Hebrew emerged (Judeo-Spanish or Ladino, Judeo-German or Yiddish, Judeo-Greek or Romaniyot, etc.).

4.3.2 The Biblical Hebrew-origin hypothesis

The article could have penetrated into European languages during the translation of the Old Testament into national languages. However, the chronology of the Bible translation, of the Old Testament in particular, shows that the article was established in Indo-European languages long before the first translations of the Old Testament.

For example, one of the first known translations of the Old Testament was made in 930-960. It was a retelling of several books of the Old Testament into English (examples 5-6) called "Caedmon manuscript".

> *(5) æfter þam wordum werod wæs on salum*
> *(After the sweet word was dark)*
> *þam - definite article, neuter, singular, dative case*
> *"Caedmon manuscript", 930-960*

> *(6) Swa him mihtig god*
> *þæs dægweorces deop lean forgeald*
> *(the mighty God recompensed to him a high reward for that day's work)*
> *þæs - definite article, masculine, singular, genitive case*
> *"Caedmon manuscript", 930-960*

But the literary monuments written a century before this translation show that the English language had already acquired the article by the time the Old Testament was translated from Old Hebrew in 10th century:

> *(7) Her Hengest & Horsa fuhton wiþ Wyrtgeorne þam cyninge, in þære stowe þe is gecueden Agælesþrep*

(This year Hengest and Horsa fought with Wurtgern the king on the spot that is called Aylesford)
þam – definite article, masculine, singular, dative case
þære – definite article, feminine, singular, dative case
"The Anglo-Saxon Chronicle", 9th c.

As follows from the chronology of literary monuments, the translation of the Old Testament from Old Hebrew did not influence the appearance of the article in Indo-European languages of Western Europe, as it was made over 100 years later.

4.4   Biblical origin

The Bible could have been a source for borrowing of the article into national languages. It consists of the Old and the New Testament. The Old Testament was written in Old Hebrew in the 13th- 1th centuries B.C. and in 132 B.C. it was translated into Koine Greek ("Septuagint"). The New Testament was written in Koine Greek in the 1st century A.D, and in 405 it was translated into Latin. Later the Latin Bible received the name "Vulgate".

The role of the church in the Early Middle Ages was very significant. After the Roman Empire had lost its power and fallen into decay, the church remained the only social institution that was common for all Western European countries. The church was not only a dominating political institution, but it was able to influence the conscience of the population. People's lives were difficult, they knew little about the world around them, and the church offered them knowledge about the organization of the world, its rules and powers. This image, based completely on the interpretation of the Bible, completely defined the mentality of the citizens. And numerous cloisters became centers of education and culture.

Since it played such an important role in the development of the Medieval culture, the Bible could have influenced the languages spoken on the territory of West Europe. But having been written in Latin, the Vulgate could not be the source of borrowing of the article, which means that the articles could have penetrated into Indo-European languages after the Bible had been translated into national languages.

The first English translations of the Bible appeared in the "Caedmon Manuscript" and the "Wessex Gospels", in the 10th century. As for the other languages in question, the Bible was translated into them a few centuries later. German had articles as early as in the 8th century (example 8), while the Bible was translated into German only in 1389; the earliest French literary monuments give examples of the use of the article in the 9th century (example 9), and the Bible was translated into French in 1297; the first complete translation of Bible into English was made in 1382, and the articles already existed in the language in the 9th century (example 10); the Spanish Bible appeared in 1280, while the language had a system of articles a century before this translation (example 11).

(8) In daz Grimensol, … ze demo Geruuinesrode, … ze dero haganinun huliu, danan in den ostaron Egelseo, dar der spirboum stuont
(to the water hole to the clearing of Gerwin to the thorn-covered swamp, then to the eastern blood Egel Sea where the rowan tree stood)
daz – definite article, neuter, singular, accusative case
demo - definite article, neuter, singular, dative case
dero - definite article, plural, genitive case
den - definite article, masculine, singular, accusative case
"Hammelburger Markbeschreibung", 777-790

(9) Voldrent la veintre li Deo inimi
(The God's enemies wanted to conquer her)
li – definite article, plural
"Cantilene de Sainte Eulalie", 880

(10) Boitius se hæle hatte se þone hlisan geþah
(The man had the name Boethius)
se – definite article, masculine, singular, nominative case
"Metres of Boethius", 9th c.

(11) Que es senior de todo el mundo
(Who is the lord of the whole world)
el - definite article, masculine, singular

"Auto de los Reyes Magos", sup. 1170

Nevertheless, there remains one translation of the Bible, which could have brought the article into Western European languages: Ulfilas' translation into Gothic.

The Gothic people arrived in Europe from Scandinavia (Gibbon, 1930), and gradually conquered the whole Iberian Peninsula, penetrated into the Roman Empire and became one of the most powerful peoples in Europe.

The Goths came from the territory of modern Sweden and Gotland Island. Soon they crossed the Baltic Sea and in the 2nd century AD they occupied the lower reaches of the Vistula River. By 230 AD they divided into Visigoths and Ostrogoths, and around the same time clashed with the Romans for the first time and forced them out of Dacia.

In 257 Ostrogoths destroyed the Scythian Kingdom and reached Eastern Crimea. During the following ten years they attacked Thrace and went as far as to Corinth and Athens. In 375 Huns defeated Ostrogoths near the Black Sea.

In 451 Ostrogoths made a military agreement with three other tribes to attack the Huns. In 488 they moved to Italy. The area defended itself for five years, until in 493 a peace treaty was signed, according to which the Gothic king and the Roman alderman ruled together. Soon Theodoric, the Gothic king, killed the Italian ruler and became the only regent, although for the Italian people he remained a Gothic chieftain and a deputy of the Roman Emperor. Theodoric was an adherent of a peaceful international policy, he wanted to blend Romans and Goths into one people, adopt the Roman culture and conquer the warlike Barbarian tribes. But at that period Romans were Catholic, while Goths were Arian (Williams, 2002: 98), which led to constant clashes between them, so Theodoric soon began to pursue Catholics.

As for the second branch of the Goths, the Visigoths, they invaded the Roman Empire in 256 and for fifteen years possessed Illyria and Macedonia. In 270 the Romans left Dacia, and the Visigoths settled there.

In 322 the Roman Empire offered the Visigoths a peace treaty: the Visigoths received the status of an ally, they provided the Roman Empire with warriors and defended the borders.

In 376 the Visigoths settled in Thrace, but the Roman aldermen constantly detained food, or exaggerated prices, which finally led to clashes. A year later, in 377 the clashes turned into an open rebellion; the Visigoths began to ravage and sack Roman territories. The Roman Empire's attempt to stifle the rebellion failed, and on 10 August, 378 the Romans were defeated. The Emperor was killed, and the remainder of the army fled. This battle played the key role in the fall of the Roman Empire, because since that time its army was no longer considered undefeatable, and the northern border was now open.

The Visigoths were ravaging the Roman settlements until 382, when the new Roman Emperor offered them a new peace treaty, the main point of which repeated the previous agreement. The treaty was cancelled when the Roman Emperor died. As a result, the Visigoths besieged Rome. The city capitulated, offering the Visigoths a huge payoff. Nevertheless, the Emperor rejected the Goths' demand for new territories, so a new besiege followed. On 24 August, 410 the Visigoths entered the city. Despite their reputation as ruthless warriors, they were merciful to peaceful citizens and did not make any serious damage to the city itself.

Two years later, when the new Visigothic king failed to gain a foothold in Italy, the Visigoths left for Gaul. During the next few years they fought as Roman allies and in 418 the Roman Empire granted them the status of confederates and they were given vast territories in Gallia Aquitania, where the Goths founded the Kingdom of Toulouse and steadily extended its borders, occupying South and Middle Gaul and almost the whole of Spain.

In 475 the Visigoths wrote their first legal code, which received the name "The Visigothic Code" (King, 1980). Gothic is the only East Germanic language that has a significant text corpus. Besides the Bible, there is another literary monument in Gothic, "Skeireins", a commentary on the Gospel of John, which was originally composed in Gothic. The volume of the surviving texts makes it possible to reconstruct, though not completely, the grammatical structure of the Gothic language.

Gothic was a Germanic languages with a lot of archaic features of Indo-European, i.e. a rich system of declensions (nominative, genitive, dative, accusative and vocative cases) and three genders (feminine, masculine and neuter). By the 9th century it went extinct. The language had already been in decline since the middle of the 6th century, due to the military defeats of the Visigoths and their conversion to Catholicism and adoption of Latin as a church language.

The Gothic verb had two tenses (present and preterite), three moods (indicative, subjunctive and imperative), two voices (active and medial) and three numbers (singular, dual and plural). All verbs were also divided into two types according to the conjugation: thematic (characterized by a thematic vowel added between the root of the verbs and the inflexional suffix) and athematic (in this case the suffix was added directly to the root). It is noteworthy that both types existed both in Old Greek and Latin. Nouns in Gothic were divided into a large variety of declensions according to the form of the stem.

All major types of pronoun were also present in Gothic: personal, possessive, interrogative, indefinite, relative and

demonstrative. All these pronouns were could be inflected; the patterns were similar to those of the nouns. A simple demonstrative pronoun was also used as the definite article.

The three forms of the Gothic definite article (*sa, so* and *þata* for masculine, feminine and neuter respectively) are derived from the Proto-Indo-European roots *so, *tod and *seh (Wright, 1910). And they are cognate with the Greek definite article *ó, τó, ή*.

The Bible translation, which is the first and the most complete surviving literary monument in the Gothic language, was produced around 350 by Ulfilas. A bishop and missionary, non-Gothic by origin, he was enslaved by Goths when he was born or when he was young, so he was raised as a Goth.

Ulfilas is considered to be the inventor of the Gothic written language, which had not existed until the Bible's translation. Some researches claim that the alphabet was derived from that of Koine Greek, while others argue that some of the Gothic letters have Runic or Latin origin.

Besides the alphabet, Ulfilas not only borrowed a big number of Greek words and usages, he also often copied the syntax from the original text, so sometimes the Gothic Bible resembles an interlinear translation of Koine Greek (Falluomini, 2005). The surviving literary monuments of the Gothic Bible present numerous examples of the use of articles as a separate grammatical category (examples 12-13). Moreover, the surviving text of "Skeireins", which was originally written in Gothic by a native speaker, also contains examples of article use (examples 14-15).

> *(12) jah qaþ Zakarias du þamma aggilau*
> *(And Zacharias said unto the angel)*
> *þamma - definite article, masculine, singular, dative case*
> *"The Gothic Bible", ~350*

> *(13) iþ Iesus qaþ du imma: laistei afar mis jah let þans dauþans <ga>filhan seinans dauþans*
> *(But Jesus said unto him, Follow me; and let the dead bury their dead)*
> *þans - definite article, masculine, plural, accusative case*
> *"The Gothic Bible", ~350*

> *(14) þizos manasedais gawaurhtedi uslunein*
> *(might accomplish the redemption of the world)*
> *þizos - definite article, feminine, singular, genitive case*
> *"Skeireins", date unknown*

> *(15) ei galaisjaina sik bi þamma twa andwairþja attins jah sunaus andhaitan*
> *(They should learn to acknowledge the double personality of the Father and the Son)*
> *þamma - definite article, neuter, singular, dative case*
> *"Skeireins", date unknown*

There is a significant correlation between the area of the article spread and the territory of Gothic conquests. Figure 4 shows the spread of the article in Indo-European languages spoken on the territory of modern Europe and Figure 5 shows the map of Europe by the 6th century. Peoples inhabiting Europe lived in close social, economic and political contact with each other, thus the probability of borrowing into the languages as a result of areal contacts was very high.
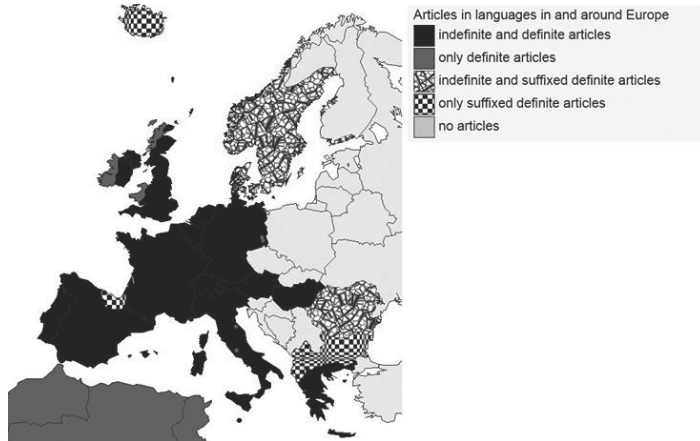
Figure 4. Distribution of definite articles in Europe. Available at http://en.wikipedia.org/wiki/Article_(grammar)#media viewer/ File:EuropeArticleLanguages.png (17 August, 2014). The picture was optimized for black-and-white format

Figure 5. Map of Europe at the end of the 6th century (Shepherd, 1923-36)

Thus, based on grammatical structures of the languages represented by literary monuments of the early Middle Age, we can say that the article as a grammatical category appeared in Old English, Old German, Old Spanish and Old French before the 7th century. It corresponds to the period of wide propagation of the Gothic Bible, which, along with the Vulgate, played an important role in the education of the elites and in the formation of the linguistic standard. As witnessed by the literary monuments, the Gothic Bible contained the article, which Ulfilas probably borrowed from Koine Greek as a grammatical category necessary for the most precise translation of the origin, i.e. the New Testament in Koine Greek. Later on, the Gothic Bible could have exerted considerable influence on the translations of the Bible into other national languages. Apparently, by the time these translations were made, the article in Indo-European languages of Western Europe had already become a linguistic standard. It is attested in numerous literary monuments (see examples: for

English 16, for French 17, for German 18, for Spanish 19). As for Portuguese, the first extant literary monuments, written in Galician-Portuguese, date back to the 13ᵗʰ century, while numerous other religious texts were completely destroyed during the inquisition.

(16) Her Iohannes se godspellere in Pathma þam ealonde wrat þa boc Apocalipsis
(This year John the evangelist in the island Patmos wrote the book called "The Apocalypse")
se - definite article, masculine, singular, nominative case
þam – definite article, neuter, singular, dative case
þa – definite article, feminine, singular, accusative case
"The Anglo-Saxon Chronicle", 9th c.

(17) Ne sai le lueu ne ne sai la contrede
(I know neither the place nor the country)
le – definite article, masculine, oblique case
la – definite article, feminine, singular
"La vie de Saint Alexis", 1040

(18) Araugit ist in des aldin uuizssodes boohhum
(It is revealed in the books of the Old Testament…)
des – definite article, masculine, singular, genitive case
"Althochdeutscher Isidor", 790

(19) Nacido es el Criador
que de las gentes es senior
(Nacido is the God, which is the master of teeth)
el - definite article, masculine, singular
las - definite article, plural
"Auto de los Reyes Magos", sup. 1170

5. Discussion

One of the most significant works dedicated to the spontaneity hypothesis of the development of the article in Indo-European languages is Bauer (2007).

The author claims that the article appeared in Romance languages during the transition from Latin and became a result of the process of the increasing use of demonstratives, while the non-Romance languages, which acquired the article as a grammatical category at the same time period, are not taken into consideration. Bauer concedes that Koine Greek could have exerted a certain influence on Latin/Romance, but it was neither the source nor the incentive of the development of the article in the Romance languages. Probably, the reason of this statement was the absence of online resources of the Gothic languages.

As we have shown in the present study, Ulfilas' Bible was the earliest translation of the New Testament into a national language, and the translations into English, German, French and Spanish appeared a few centuries later. We claim that the apparently simultaneity of the development of the article in English, German, French and Spanish was a result of the period of Gothic conquests (from 350, when the Gothic Bible appeared, till the beginning of the VI century). The existence of its unrelated demonstratives apparently developing into definite articles in each of the languages mentioned above has been used as an argument in favor of the independency of the development (Greenberg, 1978). Inasmuch as we agree that the forms developed independently we do not contradict Greenberg's theory, but we believe that the advent of the Gothic Bible became the impetus to the borrowing of the article as a matter (Sakel, 2007).

Having specified the date and the source of borrowing of the article in the wide range of Indo-European languages in Western Europe we can make a number of new assumptions concerning the development of such grammatical categories as word order, case system, inflection, topic-comment, prosodic stressing of definiteness/indefiniteness. This question requires detailed study. It concerns the Indo-European languages that have undergone articlization (authors' term: separating of the article as an independent part of speech), as well as the languages that are in long linguistic contacts with them (e.g., Uralic).

Evidently, the influence of the Gothic Bible on the formation of other grammatical categories of Indo-European languages is also of great interest. There are likely other grammatical categories, not connected with the expression of definiteness/indefiniteness, whose formation could have been influenced by the Gothic Bible.

Summing up our arguments, the comparison of the map of the article spread in Europe and the map of Gothic

conquests is in strong support of our assumption. Another argument is that the New Testament written in Koine Greek had the article as a grammatical category, which penetrated into Gothic when Ulfilas translated the New Testament in the middle of the IV century. Third, the Gothic Bible appeared long before the first translations of the Bible into French, German, English and Spanish and was followed by the period of a strong Gothic rule in Western Europe, which means that Gothic could exert a significant influence on the languages spoken on this territory at that time. Thus, we have three arguments in favor of the suggested hypothesis. Separately, each of them, undoubtedly, is not sufficient, but the concurrence of the three arguments provides a strong basis of the suggested hypothesis.

Due to the complete absence of literary monuments in English, French, German or Spanish until the end of the VIII century we cannot provide the direct proof of our hypothesis, but nevertheless consider it a viable hypothesis worth considering along with the other hypotheses discussed.

6. Conclusion

This paper has discussed the main function of the article and its connection with other grammatical categories. Several hypotheses concerning the origin of the article in Indo-European languages of Western Europe were studied, including: a hypothesis about spontaneous and independent development of the article in several languages and language families, the penetration of the article from Arabic, from Hebrew, and the borrowing of the article due to translations of the Bible into national languages. The historical analysis of literary monuments in Old French, Old Spanish, Old German and Old English showed that the languages in question had acquired the article long before the Bible was translated into them. Basing on the chronology of literary monuments in ancient European languages and on the comparison of the map of article spread in European languages and of the map of Gothic conquests in the 6th century AD, we suggested the hypothesis that Ulfilas' Gothic Bible became the source of borrowing of the article in Western Indo-European languages.

7. Acknowledgments

References

Anisimov, Ivan, Polyakov, Vladimir & Solovyev, Valery. 2013. *Database "Languages of the World". New Version. New Research Horizons.* Collection of Papers of the First International Forum on Cognitive Modeling (14-21 September, 2013, Italy, Milano-Marittima). In 2 parts. / Edited by S. Masalóva V. Solovyev. - Part 1. Cognitive Modeling in Linguistics: Proceedings of the XIV International Conference « Cognitive Modeling in Linguistics. CML-2013». Rostov-on-Don: Southern Federal University Press, 2013. P. 27-34. ISBN 918-5-87872-731-0

Bauer, Brigitte L.M. 2007. *The definite article in Indo-European. Emergence of a new grammatical category?* Studies in Language Companion Series (SLCS). V. 89. Nominal determination. Typology, context constrains and historical emergence. Edited by Elisabeth Stark, Elisabeth Leiss and Werner Abraham. Amsterdam/Philadelphia: John Benjamins Publishing Company. pp. 103-139.

Bisle-Müller, Hansjörg. 1991. A*rtikelwörter im Deutschen. Semantische und pragmatische Aspekte ihrer Verwendung*. Niemeyer: Tübingen. [in German]

Brachet, Auguste. 1876. *Dictionnaire étymologique de la langue française*. Paris: J. Hetzel. http://www.archive.org/details/Dictionnairety mologiqueDeLaLangueFrancaise

(Accessed on 12 August, 2014) [in French]

Collins, Roger. 1989. *The Arab Conquest of Spain 710-797*. Oxford, UK/Cambridge, USA: Blackwell.

Dryer, Matthew S. 2013. Definite articles. In: Dryer, Matthew S. & Martin Haspelmath (eds.). *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. http://wals.info/chapter/37 (accessed 17 August, 2014)

Falluomini, Carla. 2005. *Textkritische Anmerkungen zur gotischen Bibel*. AnnalSS 5. 311–320. [in German]

Furst, Julius. 1849. The Jews in Spain under the Visigoths. *The Occident and American Jewish Advocate*. Electronic archive. http://www.jewish-history.com/occident/volume7/nov1849/visigoths.html (Accessed on 11 August, 2014)

Gibbon, Edward. 1930. *The Decline and Fall of the Roman Empire*. Plain Label Books.

Greenberg, Joseph H. 1978 (reprinted in 2004). *Genetic Linguistics: Essays on Theory and Method.* Oxford: Oxford University Press. 460.

Hoad, Terry F.(ed.). 1996. *The Concise Oxford Dictionary of English Etymology*. http://www.oxfordreference.com/view/10.1093/ acref/9780192830982.001.0001/acref-9780192830982 (Accessed on 11 August, 2014)

King, P. D. 1980. Chindasvind and the first territorial law-code of the Visiogothic kingdom. In James, Edward (ed.). *Visigothic Spain: New Approaches*. 131–157. Oxford: Clarendon Pess.

Lapesa, Raphael. 1942. *Historia de la lengua española.* Madrid: Escelicer. [in Spanish]

Lomax, Derek W. 1978. *The Reconquest of Spain*. London/ New-York: Longman.

Polyakov, Vladimir, Solovyev, Valery, Wichmann, Søren & Oleg Belyaev. 2009. Using WALS and Jazyki Mira. *Linguistic Typology* 13. 135-165.

Quintana, Lucía & Juan Pablo Mora. 2002. *Enseñanza del acervo léxico árabe de la lengua española*. ASELE. Actas XIII. 705. [in Spanish]

Roth, Norman. 1994. *Jews, Visigoths and Muslims in Medieval Spain: Cooperation and Conflict*. Leiden/New York: Brill.

Ryding, Karin C. 2005. *A Reference Grammar of Modern Standard Arabic* (6th ed.). Cambridge: Cambridge University Press.

Sáenz-Badillos, Angel. 1993. *A History of the Hebrew Language*. Cambridge: Cambridge University Press.

Sakel, Jeanette. 2007. *Grammatical borrowing in cross-linguistic perspective*. Berlin: Mouton de Gruyter.

Shepherd, William. 1923-26. Historical Atlas. University of Texas at Austin. http://www.emersonkent.com/map_archive/germanic_roman_526.htm (Accessed on 10 August, 2014)

V.D. Solovyev, V.N. Polyakov. 2013. *Database "Languages of the World" and its application. State of the art*. Computational Linguistics and Intellectual Technologies. Issue 12 (19). v. 1. Papers from Annual Int. Conf. DIALOG-2013. p. 672-681.

Williams, Rowan. 2002. Arius: Heresy and Tradition. Wm. B. Eerdmans Publishing Co.

Wright, Joseph. 1910. *Grammar of the Gothic Language*. Oxford: Clarendon Press.

Online Sources of Information

http://asjp.clld.org/ (accessed 2 February, 2014)

http://en.wikipedia.org (accessed 23 August, 2014)

http://wals.info (accessed 24 August, 2014)

www.archive.org (accessed 24 August, 2014)

www.asc.jebbo.co.uk (accessed 24 August, 2014)

www.bible-researcher.com (accessed 24 August, 2014)

www.dreamofrood.co.uk (accessed 24 August, 2014)

www.fordham.edu (accessed 24 August, 2014)

www.heorot.dk (accessed 24 August, 2014)

www.hs-augsburg.de (accessed 24 August, 2014) [in German]

www.linguistics.ruhr-uni-bochum.de (accessed 24 August, 2014) [in German]

www.ru.scribd.com (accessed 24 August, 2014)

www.russianplanet.ru (accessed 24 August, 2014) [in Russian]

www.sacred-texts.com (accessed 24 August, 2014)

www.studylight.org (accessed 24 August, 2014)

www.wikipedia.org (accessed 24 August, 2014)

www.wikisource.org (accessed 24 August, 2014)

www.wordhord.org (accessed 24 August, 2014)

www.wulfila.be (accessed 24 August, 2014)

https://cloud.mail.ru/public/eaa329cb8c0b/DBlang/ (The database "Languages of the World" available for download, accessed 17 October, 2015)