Automatic Translation of Tlemcen's Named Entities (From Arabic to English)

Mohammed Dib .Maitre assistant University of Mascara
Department of English language

Abstract:

By definition all cognitive experience and its classification is conveyable in any existing language, and despite the non existence of some terms in the different culture, this would not stop translators to express them in the target language, whether manually or automatically. Yet the overriding thing to mention is that in addition to poetry which is claimed to be untranslatable, since the form of the verse contributes to build the meaning of the text, anthroponyms or proper nouns, however, are very difficult to translate and have a special translation technique. This research is an attempt to translate the place names of Tlemcen using Buck Walter transliteration.

Introduction:

This paper presents a study of automatic translation of Tlemcen's toponyms which falls in the field of toponymy .This latter studies the place names and encompasses a lot of types such as: hydronyms (names of bodies of water), oronyms (names of mountains) geonyms (general names for streets) and oeconyms (names of populated places). In fact automatic translation of anthropnyms is a challenging task because it requires a special technique and treatment. In this work the empirical approach is adopted; one of the pivotal approaches of computational architecture, it hinges on a training phase whose role is to make transformation of an intermediary representation to another one. The first part of this paper overviews toponymy, translation and strategies, the second part deals with the alignment of a parallel corpora relying on direct translation and Buck Walter translation, and the application of competitive linking algorithm.

Toponymy

Toponymy is a Greek word composed of topos (place) and onyma (Aeolic and Doric form of onoma) name. It embraces all

aspects of place names whether theoretical or practical. [1] On the theoretical side, place names can give us tremendous information regarding the physical environment, culture, history of this space language, and even the people who lived in this place. On the practical side, however, place names deals with administrative bodies for instance geographical name organizations or commissions.[2]So place names play a significant role in shaping both the personality of individuals and strengthen the government policy. In this respect Valliéres argues that Toponymy, whether we realize it or not, is part of the primary needs of every person in our planet."[3]

Toponymic studies have classically attempted to answer the WH- questions for each place name: what is it? Where is it? Who named it? When was it named? And why was it given that name?[4]H.L. Mencken (1967 [1919]: 643) argues that toponyms fall into eight classes: From personal names, Transferred from other and older places names, Native American names, Foreign language names (e.g. Dutch, Spanish, French, German, Scandinavian) ,Biblical and mythological names, Descriptive of localities Suggested by local flora, fauna, or geology Purely fanciful names. [5] The purpose of toponymy is to establish a standardized feature catalogue to build Gazetteer code containing three different things: The identification of a set of semantic components related to topographic features e.g. AF and ANCH , the production of each feature sets within the catalogue by a logical sequence of those components e.g. Aero, Agricultural Apical, ,Arboreal , Barrier, Bathymetric, and the establishment of a glossary of included feature set.eg. Abattoir, Abbey, Aboriginal outstation

Translation:

There is no doubt that any translator who wants to translate a given text, should trace a road map or follow a strategy to reach his

goal, because translation is an intensive labour, but before all let's have some definitions about translation.

Translation is the expression of meaning in a given language, the source language in terms of another, the target language.[1]It is also the act or an instance of translating a written or spoken expression of the meaning of a word, speech, book, etc…in another language[2] , that is, all cognitive experience and its classification is conveyable in any existing language, and despite the non existence of particular terms in different cultures, this does not stop translators to express them in other ways in the target language. R.Jakobson claims that only poetry by definition is untranslatable, because the form of the words in the verse governs the meaning of the text.[3]Translation can be seen as a process and a product at the same time, the first sense focuses on the rendering of source text into the target text. The second sense however, deals with the concrete product of the translator. [4]Regarding the types of translation, Roman Jakobson divided translation into three types; intralingual translation, interlingual translation and intersemiotic translation.

Intralingual translation: Translation here is made within the same language and relies upon rewording and paraphrase. e.g. translation of dialect

Interlingual translation: Translation from one language to another.

Intersemiotic translation: Translation of the verbal sign by a non-verbal sign, for example music or image.[5]

Strategy:

Strategy is a specific method to approach a problem or task, modes or operation for achieving a particular end, planned designs for controlling and manipulating certain information, and a number of possible ways to solve a problem.[6]

Translations strategies:

There are four strategies in translation they are as follows: transference strategy, transliteration strategy per se, and combined strategy.

1. transference strategy, i.e. do-not translate;
2. transliteration strategy, i.e. phonetic or spelling rendering;
3. translation strategy per se, i.e. do translate
4. combined strategy, i.e. applying more than one from the above mentioned three

strategies.[1]

The first type of translation (The transference strategy) is usually applied for toponyms which do not need any rendering, therefore are left not translated e.g. organization
names (Babych and Hartley, 2003). The second type of translation, however, (transliteration strategy) is used between languages that do not share the same alphabet, therefore uses a rendering of a source language word into a target language word throughout its phonemic representation i.e., grapheme-phoneme-grapheme rendering.

Automatic translation:

It is a process and a product that marry computing power and computerized analysis of language to the human's ability to analyze sense and determine appropriate forms in the other language.[2]. It is useful for the natural language processing and is designed mainly for automatic processing because it uses the standard *ASCII characters, encoded on 7bits and it are reversible between Arabic and ASCII characters.[3]

---

[1] English-Latvian Toponym Processing:Translation Strategies and Linguistic Pattern2009 European Association for Machine Translation.Proceedings of the 13th Annual Conference of the EAMT, pages 81–87,Barcelona, May 2009

*** American Standard Code for Information Interchange(Microsoft® Encarta® 2009. © 1993-2008 Microsoft Corporation.)**

Empirical approach and Buck Walter transliteration[1] are used in our work to show this combination. The former means to select nouns and put them in a given context and little by little try to adjust the text. The latter, however, relies on transcription which adapts the phonetic image of a foreign toponym to the graphical norm and to each phoneme corresponding to one or more signs . The principal advantage of transcription is to be embedded in the target language, therefore allows a great legibility and a so called pronunciation.[2]

Since we are dealing with the translation of anthroponyms, we have to mention some kinds of ambiguity.

Toponyms are ambiguous therefore Leidner 2007 argues that there are 3 kinds of toponymical ambiguity

- Morpho syntactic ambiguity: A word can be a toponym and a common noun; Ewe a West African language and ewe female of a sheep, Hook a populated place in the UK and hook a common noun.
- Referential ambiguity one toponym may refer to more than one place of the same type e.g. Sidi said in Tlemcen , Sidi said in Temouchent and Sidi said in Mascara, or Riga a populated place and capital of Latvia and populated place in USA.
  Nida argues that there are two types of meaning; referential meaning and connotative meaning. Referential meaning or denotation which stands for the signs or symbols of the word,by contrast connotative meaning refers to the emotional reaction created in the reader by a word.[3]

[1] Transliteration : The word transliteration is taken from the latin word littera , letter it stands for letter for letter substitution. It is a method of names conversion between different scripts and not languages.ie each character in the source script is replaced by a corresponding character in the target script.

***ASCII American Standard Code for information Interchange**

- Feature type ambiguity one toponym may refer to more than one place of different type
- Eponymical ambiguity that is, names of people are named after people or deities.[1]

Buckwalter Transliteration

Arabic letters	**Time New Romans**	**Buck Walter symbols**
ا	A	lone hamza: '
إ ب	b	hamza on alif: >
إ ت	t	hamza below alif: <
إ ث	v	hamza on wa: &
ج	j	**hamza on ya: }**
ح	H	alif : `
خ	x	madda on alif: |
د	d	alif al-wasla: {
ذ	*	dagger alif :`
ر	r	alif maqsura: Y
ز	z	harakat
س	s	fatha: a
ش	$	damma: u
ص	S	kasra: i
ض	D	fathatayn : F
ط	T	dammatayn: N
ظ	Z	kasratayn K
ع	E	shadda:

		~
غ	g	sukun: o
ف	f	ta marbouta: p
ق	q	tatwil: _
ك	k	
ل	l	
م	m	
ن	n	
ه	h	
و	w	
ي	y	
ى / ي[11]		

Practical survey

English and Arabic have different Alphabets, and in order to find equivalents between toponyms or anthroponyms in bilingual parallel texts, we find it a must to rely upon phonetic similarities shared by some units, there fore use Buck Walter transliteration to be able to do texts alignment, because it indicates the existing correspondences between words of the source text and those of the target one, therefore defines for each source word the target word which translates.

The corpus used in this work contains some anthroponyms of Tlemcen put into an Arabic context, then translated into English, and then segmented into phrases and words. In order to make alignment between the two corpora, transcription using Buck Walter transliteration is made, and to calculate the number of occurrences of the source and the target words, the dice coefficient is applied to build a matrix. The word in the source text is represented by "e", while the word in the target text is represented by "f". In doing so, we can calculate how many times is the word "e" aligned with the word "f" in both the source text and the target text.

Here is the dice coefficient

$$2.c(e,f)$$

dice(e,f)=……..

$$\frac{C(e)+c(f)}{}$$

e and f are words in source language and target language

Tlemcen's anthroponyms

Arabic Written form	Times new Roman	Buck Walter transliteration
لالا ستي	Lala setti	lala| satiy
باب الجياد	Babaljiayad	ba|b{lji ya|d
باب الحديد	Bab Alhdid	ba|b{lhodiyd
باب الخميس	Babalkhmis	ba|b{l xomiys
باب القرمادين	Babalkarmadine	ba|b{lqarmadiyn
باب وهران	Babwahran	ba|bwahra|n
سيدي الحلوي	Sidialhalloui	sidi{lhalwiy
سيدي بومدين	Sidiboumedienne	sidibu madyan
المشور	Almashwar	lma$ w a r
المنصورة	Mansourah	lmansurah
جامع سيدي ابرهيم	SidiBrahim Mosque	sidibrahiym
عرست ديدو	Arsat didou	Earsat diyduw

Parallel corpus

اختيرت عاصمة للثقافة الاسلامية نظرا للحضارات التي تلمسان عاصمة الثقافة الإسلامية. و تتميز بمناخ بارد في فصل الشتاء و حار في فصل شهدها. تقع تلمسان في غرب الجزائر الصيف.تحيط مدينة تلمسان جبال عديدة مثل جبل لالاستي و جبل سيدي الطاهر و جبل لوريت .توجد في مدينة تلمسان سبعة أبواب نحو باب الجياد و باب الحديد و باب وهران

باب القرمادين وباب الخميس باب زير و يوجد في حي باب الجياد عرست ديدو.تشمل و تلمسان عدة مساجد مثل

الجامع الكبير و جامع سيدي ابر هيم و سيدي الحلوي و سيدي بومدين و مسجد مولاي عبد الله الشريف التلمساني .

فيما يخص الآثار الموجودة في منطقة تلمسان فهناك المنصورة و المشور و باب القرمادين.

Tlemcen, the capital of Islamic culture. It was chosen as a capital of Islamic culture due to the civilizations that has witnessed .Tlemcen is located in the west of Algeria and it is characterized by a cold weather in winter and hot one in Summer.Tlemcen is surrounded by many mountains ,like Lala setti ,sidi Tahar and Lourit.There are in Tlemcen seven doors ,like Babaljiad, Babalhadid,Babwahran Bab alkarmadine,balkhamis,Babzir and there is in Babaljiad Arsat didou.

Tlemcen contains many mosques, like big mosque Sidi brahim mosque,Sidi alhalloui, sidi boummedienne and Moulay abdellah cherif atilimssani.

As for monument found in Tlemcen, there is almansura almashwar,wa babalkarmadine

Segmentation of Corpus in phrases

1. تلمسان عاصمة الثقافة الإسلامية.
2. اختيرت عاصمة للثقافة الإسلامية .
3. نظرا للحضارات التي شهدتها
4. تقع تلمسان في غرب الجزائر
5. وتتميز بمناخ بارد في فصل الشتاء
6. و حار في فصل الصيف
7. تحيط مدينة تلمسان جبال عديدة

مثل جبل لالاستي و جبل سيدي الطاهر و جبل لوريت .8

.توجد في مدينة تلمسان سبعة أبواب .9

باب القرمادين نحو باب الجياد و باب الحديد و باب وهران و.10
وباب الخميس باب زير

و توجد في حي باب الجياد عرست ديدو.11

تشمل تلمسان عدة مساجد مثل.12

الجامع الكبير و جامع سيدي ابر هيم و سيدي الحلوي و سيدي بومدين و .13
مسجد مولاي عبد الله الشريف التلمساني

فيما يخص الاثارا الموجودة في منطقة تلمسان **14.**

فهناك المنصورة و المشور و باب القرمادين. **15.**

1. **Tlemcen the capital of Islamic culture**
2. **It was chosen as a capital of Islamic culture due to the civilizations that has witnessed**
3. **And it is located in the west of Algeria**
4. **It is characterized by a cold weather in winter**
5. **and hot one in Summer**
6. **Tlemcen is surrounded by many mountains**
7. **like Lala setti mountain sidi Tahar mountain and Lourit**
8. **There are in Tlemcen seven doors**
9. **like Babaljiad, Babalhadid,Babwahran Bab alkarmadine,balkhamis,Babzir**
10. **and there is in Babaljiad Arsat didou**
11. **Tlemcen contains many mosques**
12. **like the big mosque Sidi brahim mosque,Sidi alhalloui,**
13. **sidi boummedienne and Moulay abdellah cherif atilimssani**

14. **As for monuments found in Tlemcen,**

15. there is almansura almashwar,wa babalkarmadine

Alignment of texts into phrases

tilamsa|n Ea|Simap {vaqafa {l<isla|miyah

Tlemcen the capital of Islamic culture

>u xtyrat Ea|SimatF li {vaqafa p l<isla|miyah

It was choosen as a capital of Islamic culture

na ZarF lil HaDa|ra|t {laity $ahidatha|

taqaE fiy garbi {ljaza|**}r**

It is located in the west of Algeria

tatamay~azu bi muna|x K ba|ridK fiy faSli{$ita|'i

It is characterized by a cold weather in winter

Wa ha|rK fiy faSli{$ayfi

and a hot one in Summer

tuhiyTu madiyn p tilamsa|n Eid~atu jiba|l

Tlemcen is surrounded town by many mountains

mivla lala| satiy jabal sidi{Tahar wa lurit

like Lala setti mountain sidi Thar mountain and Lourit

Tuwjadu fiy tilamsa|n sabEa pu >abwa|b

There are in Tlemcen seven doors

mivla ba|b{lji ya|d ba|b{lhdiyd ba|bwahra|n

like Babaljiad, Babalhadid, Babwahran

ba|b{lqarmadiyn ba|b{l xomiys

Bab alkarmadine balkhamis,Babzir

Wa yuwjad fiy ba|b{lji ya|d Earsat diyduw

and there is in Babaljiad Arsat didou

tilamsa|n tamulu Eid~atu masa|jid

Tlemcen contains many mosques

mivla kabiyr masjid sidibrahiym sidi{lhalwiy

like big mosque Sidi brahim Sidi alhalloui,

sidibumadyan wa muwla|y Eabdalah $o riyf {atilamsa|ni

sidi boummedienne and Moulay abdellah cherif atilimssani

Implementation of alignment model by words

tilamsa|n Ea|Simap {vaqafa {lisla|miyah

Tlemcen the capital of Islamic culture

dice(tilamsa|n, Tlemcen) dice (tilamsa|n, the) dice(tilamsa|n, capital) dice (tilamsa|n, of)

dice(tilamsa|n, Islamic) dice (tilamsa|n, culture)

dice(Ea|Simap, Tlemcen) dice(Ea|Simap, the)dice(Ea|Simap ,capital)dice(Ea|Simap, of) dice(Ea|Simap, Islamic) dice(Ea|Simap, culture)

dice{vaqafa, Tlemcen) dice{vaqafa, the) dice{vaqafa, capital) dice{vaqafa, of) dice{vaqafa, Islamic) dice{vaqafa, culture)

dice ({lisla|miyah,Tlemcen) dice ({lisla|miyah,the) dice ({lisla|miyah,capital) dice ({lisla|miyah,of) dice ({lisla|miyah,Islamic) dice ({lisla|miyah,culture)

dice(tilamsa|n)=7

dice(Tlemcen)=7

dice(tilamsa|n ,Tlemcen) =2c((tilamsa|n ,Tlemcen)/c(tilamsa|n)+ c(Tlemcen)

tilamsa|n is aligned 5times with Tlemcen so we get the following:

dice(7 ,7)= 2c(7 ,7) /c(7)+ c(7)this means 2(7 ,7)/7+7= 14/14=1

dice (tilamsa|n)=7

dice (the)=3

dice (tilamsa|n,the)=dice(7,3). dice(7,3)= 2c(7 ,3) /c(7)+ c(2)=2x3/7=0.857

dice(tilamsa|n)7

dice (capita)l=2

dice(tilamsa|n ,capital)= dice(7,2) this means dice(7,2)= 2c(7 ,2) /c(7)+ c(2)= 0.444

dice (tilamsa|n) =7

dice (of)=2

dice (tilamsa|n, of) =dice(7,2). dice(7,2)= 2c(7 ,2) /c(7)+ c(2)=2x2/7=0.571

dice(tilamsa|n)=7

dice(Islamic)=2

dice(tilamsa|n, Islamic)= dice(7,2) this means dice(7,2)= 2c(7 ,2) /c(7)+ c(2)= 0.444

dice (tilamsa|n)=7

dice (culture)=1

dice (tilamsa|n, culture)= dice(7,1) this means dice(7,1)= 2c(7 ,1) /c(7)+ c(1)= 0.25

dice(Ea|Simap, Tlemcen)

dice(Ea|Simap)=1

dice(Tlemcen)=7

dice(Ea|Simap, Tlemcen)= dice(1,7) this means dice(1,7)= 2c(1,7) /c(1)+ c(7)= 0.25

dice(Ea|Simap)=1

dice(, the)=3

dice(Ea|Simap, the)= dice(1,3) this means dice(1,3)= 2c(1,3) /c(1)+ c(3)= 0.666

dice(Ea|Simap)=1

dice(,capital)=2

dice(Ea|Simap ,capital) = dice(1,2) this means dice(1,2)= 2c(1,2) /c(1)+ c(2)=0.666

dice(Ea|Simap)=1

dice(of)=2

dice(Ea|Simap, of))= dice(1,2) this means dice(1,2)= 2c(1,2) /c(1)+ c(2)=0.666

dice (Ea|Simap)=1

dice (Islamic)=2

(Ea|Simap, Islamic))= dice(1,2) this means dice(1,2)= 2c(1,2) /c(1)+ c(2)=0.666

dice(Ea|Simap)=1

dice(culture)= 2

dice(Ea|Simap, culture)= dice(1,2) this means dice(1,2)= 2c(1,2) /c(1)+ c(2)=0.666

dice{vaqafa)=2

dice{ Tlemcen)=7

dice{vaqafa, Tlemcen)= dice(2,7) this means dice(2,7)= 2c(2,7) /c(2)+ c(7)=0.444

dice{vaqafa)= 2

dice{ the)= 3

dice{vaqafa, the) =dice(2,3) this means dice(2,3)= 2c(2,3) /c(2)+ c(3)=0.8

dice{vaqafa)=2

dice{ capital)=2

dice{vaqafa, Tlemcen)= dice(2,2) this means dice(2,2)= 2c(2,2) /c(2)+ c(2)=1

dice{vaqafa)=2

dice{of)=2

dice{vaqafa ,of)= dice(2,2) this means dice(2,2)= 2c(2,2) /c(2)+ c(2)=1

dice{vaqafa)=2

dice{ Islamic)=2

dice{vaqafa, Islamic)= dice(2,2) this means dice(2,2)= 2c(2,2) /c(2)+ c(2)=1

dice{vaqafa ,culture))= dice(2,2) this means dice(2,2)= 2c(2,2) /c(2)+ c(2)=1

dice ({lisla|miyah)=2

dice ({Tlemcen)=7

dice ({lisla|miyah,Tlemcen)= dice(2,7) this means dice(2,7)= 2c(2,7) /c(2)+ c(7)=0.444

dice ({lisla|miyah)=2

dice ({the)=3

dice ({lisla|miyah,the)= dice(2,3) this means dice(2,3)= 2c(2,3) /c(2)+ c(3)=0.8

dice ({lisla|miyah,)=2

dice (capital)=2

dice ({lisla|miyah,capital)= dice(2,2) this means dice(2,2)= 2c(2,2) /c(2)+ c(2)=1

dice ({lisla|miyah)=2

dice (of)=2

dice ({lisla|miyah,of)=)= dice(2,2) this means dice(2,2)= 2c(2,2) /c(2)+ c(2)=1

dice ({lisla|miyah)=2

dice (Islamic)=2

dice ({lisla|miyah,Islamic)=)= dice(2,2) this means dice(2,2)= 2c(2,2) /c(2)+ c(2)=1

matrix of the first two sentences

	tilamsa|n	Ea|Simap	{vaqafa	{lisla|miyah
Tlemcen	1	0.25	0.444	0.444
The	0.857	0.666	0.8	0.8
Capital	0.444	0.666	1	1

Of	0.571	0.666	1	1
Islamic	0.444	0.666	1	1
Culture	0.25	0.666	1	1

taqaEu fiy garbi {ljaza|}**r**

It is located in the west of Algeria

Dice(taqaEu , It) dice(taqaE , It) dice(taqaE , is) dice(taqaE, located) dice(taqaEu , in) dice(taqaEu, the) dice(taqaEu ,west) dice(taqaEu, of) dice(taqaEu, Algeria)

Dice (fiy,It) dice(fiy, is) dice(fiy, located) dice(fiy **,** in) dice(fiy, the) dice(fiy, west) dice(fiy, of) dice(fiy, Algeria)

Dice(garbi, It) dice(garbi, is) dice(garbi, located) dice(garbi, in)dice(garbi, the) dice(garbi ,west) dice(garbi, of) dice(garbi, Algeria)

Dice({ljaza|}**r,** It) Dice({ljaza|}**r,** is) Dice({ljaza|}**r,** located) Dice({ljaza|}**r,** in) Dice({ljaza|}**r,** in) Dice({ljaza|}**r,** the) Dice({ljaza|}**r,** west) Dice({ljaza|}**r,** of) Dice({ljaza|}**r,** Algeria)

matrix of the second two sentences

	taqaEu	fiy	garbi	{ljaza|}**r**
It				
Is				
Located				
In		1		
West			1	
of				
Algeria				1

tatamay~azu bi muna|x K ba|ridK fiy faSli{$ita|'i

It is characterized by a cold weather in winter

dice (tatamay~azu, It) dice (tatamay~azu, is) dice (tatamay~azu, Characterized) dice (tatamay~azu, by) dice (tatamay~azu, a) dice (tatamay~azu, cold) dice (tatamay~azu, weather) dice (tatamay~azu, in) dice (tatamay~azu, winter)

dice (bi, It) dice (bi, is) dice (bi, characterized) dice (bi, by) dice (bi, a) dice (bi, cold) dice (bi, weather) dice (bi, in) dice (bi, winter)

dice (muna|x K, It) dice (muna|x K, is) dice (muna|x K, characterized) dice (muna|x K, by) dice (muna|x K, a) dice (muna|x K, cold) dice (muna|x K, weather) dice (muna|x K, in) dice (muna|x K, winter)

dice (ba|ridK, It) dice (ba|ridK, is) dice (ba|ridK, characterized) dice (ba|ridK, by) dice (ba|ridK, a) dice (ba|ridK, cold) dice (ba|ridK, weather) dice (ba|ridK,in) dice (ba|ridK, in) dice (ba|ridK, winter)

dice (fiy, It) dice (fiy, is) dice (fiy, characterized) dice (fiy, by) dice (fiy, a) dice (fiy, cold) dice (fiy, weather) dice (fiy, in) dice (fiy, winter)

dice (faSli, It) dice (faSli, is) dice (faSli, characterized) dice (faSli, by) dice (faSli, a) dice (faSli, a) dice (faSli, cold) dice (faSli, weather) dice (faSli, in) dice (faSli, winter)

dice ({$ita|'i ,It) dice ({$ita|'i,is) dice ({$ita|'i,characterized) dice ({$ita|'i, by) dice ({$ita|'i, a) dice ({$ita|'i,cold) dice ({$ita|'i,weather) dice ({$ita|'i ,in) dice ({$ita|'i ,winter)

matrix of the third two sentences

	tatamay~ azu	b i	muna| x K	ba|rid K	fi y	faSl i	{$ita|' i
It							
Is							
characterized							
By							
A							
Cold							
weather							
In					1		
Winter							1

Wa ha|r fiy faSli{$ayfi

and hot one in summer

dice (Wa, and) dice (Wa, hot) dice (Wa, one) dice (Wa, in) dice (Wa, summer)

dice (ha|r, and) dice (ha|r, hot) dice (ha|r, one) dice (ha|r, in) dice (ha|r, summer)

dice (fiy, and) dice (fiy, hot) dice (fiy, one) dice (fiy, in) dice (fiy, summer)

dice (faSli, and) dice (faSli, hot) dice (faSli, one) dice (faSli, in) dice (faSli, summer)

dice ($ayfi , and) dice ($ayfi , hot) dice ($ayfi , one) dice ($ayfi , in) dice ($ayfi , summer)

matrix of the fourth two sentences

	Wa	ha|r	fiy	faSli	$ayfi
And	1				
Hot		1			
One					
In			1		
Summer					1

tuhiyTu madiyn p tilamsa|n Eid~atu jiba|l

Tlemcen town is surrounded by many mountains

dice (tuhiyTu, Tlemcen) dice (tuhiyTu, town) dice (tuhiyTu, is) dice (tuhiyTu, surronded) dice (tuhiyTu, by) dice (tuhiyTu, many) dice (tuhiyTu, mountains)

dice (madiyn p, Tlemcen) dice (madiyn p, town) dice (madiyn p, is) dice (madiyn p, surrounded) dice (madiyn p, by) dice (madiyn p, many) dice (madiyn p, mountains)

dice (tilamsa|n, Tlemcen) dice (tilamsa|n, is) dice (tilamsa|n, surrounded) dice (tilamsa|n, town) dice (tilamsa|n, by) dice (tilamsa|n, many) dice (tilamsa|n, mountains)

dice (Eid~atu, Tlemcen) dice (Eid~atu, is) dice (Eid~atu, surrounded) dice (Eid~atu, town) dice (Eid~atu, by) dice (Eid~atu, many) dice (Eid~atu, mountains)

dice (jiba|l, Tlemcen) dice (jiba|l, is) dice (jiba|l, surrounded) dice (jiba|l,by) dice (jiba|l,many) dice (jiba|l, mountains)

matrix of the fifth two sentences

	tuhiyTu	madiyn p	tilamsa|n	Eid~atu	jiba|l
Tlemcen			1		
town					
Is					
surrounded					
by					
Many					
mountains					1

mivla lala| satiy jabal sidi{Tahar wa lurit

like Lala setti mountain sidi Tahar mountain and Lourit

dice (mivla, like) dice (mivla, Lala) dice (mivla, setti) dice (mivla, mountain) dice (mivla, sidi) dice (mivla, Tahar) dice (mivla, mountain) dice (mivla, and) dice (mivla, Lourit)

dice (lala| , like) dice (lala| , Lala) dice (lala| , setti) dice (lala| , mountain) dice (lala| , sidi) dice (lala| , Tahar) dice (lala| , mountain) dice (lala| , and) dice (lala| , Lourit)

dice (wa,like)

dice (satiy, like) dice (satiy, Lala) dice (satiy, setti) dice (satiy, mountain) dice (satiy, sidi) dice (satiy, Tahar) dice (satiy, mountain) dice (satiy, and) dice (satiy, Lourit) dice (sidi, like) dice (sidi, Lala) dice (sidi, setti) dice (sidi, mountain) dice (sidi,

sidi) dice (sidi, Tahar) dice (sidi, mountain) dice (sidi, and) dice (sidi, Lourit)

dice ({Tahar, like) dice ({Tahar, Lala) dice ({Tahar, setti) dice ({Tahar, mountain) dice ({Tahar, and) dice ({Tahar, Lourit)

We do the same process until we finish the alignment of all phrases and words. After finishing alignment to construct matrices, we move to the application of the competitive linking algorithm

competitive linking algorithm is used to reduce translation errors and enhance the reliability of the estimation. And it determines the most probable translation pairs; the source text and the target one.[1]

How to apply the competitive linking algorithm,

We use the competitive linking to maximize the M (i,j) According to the following matrix(of the first two sentences) :

matrix of the first two sentences

	tilamsa|n	Ea|Simap	{vaqafa	{lisla|miyah
Tlemcen	1	0.25	0.444	0.444
The	0.857	0.666	0.8	0.8
Capital	0.444	0.666	1	1
Of	0.571	0.666	1	1
Islamic	0.444	0.666	1	1
Culture	0.25	0.666	1	1

Dice (tilamsa|n, Tlemcen)=1

Dice (tilamsa|n, the)= 0.857

Dice (tilamsa|n, capital)= 0.444

Dice (tilamsa|n, of)= 0.571

Dice (tilamsa|n, Islamic)= 0.444

Dice (tilamsa|n, culture)= 0.25

In the first step we align the pair of words (i,j) which maximizes M (i,j)

For the above examples, We look for the max of M (i,j) within the matrix(the greatest dice coefficient) in order. We start by the first line where we have the max of M (i,j) =1.000 ,it is the dice ((tilamsa|n, Tlemcen) and is aligned with the word Tlemcen

In the second step we delete the line "i "and the column" j "from the matrix and remove the line where there is he word tilamsa|n, and we continue to do the same process until all words are aligned.

Conclusion:

Our work shows that automatic translation of toponyms requires Buck Walter translation, because it is considered as the best method for alignment between two languages that do not share the same alphabet ,more over to acquire good results once adopting the empirical approach requires a large corpus because a small corpus would not give satisfactory results.

References

[1] Naftali Kadmon. Toponymy The Lore,Laws and Language of geographical Names..p:4

[2] Ibid
[3] Ibid
[4] Motivations for Naming :A Toponymic Typology.ANPS Technical paper No. 2 p:1
[5]ibid
[6] A standardized Geographic Feature Catalogue for toponymic research ANPS Technical Paper n 1 .2008
[7] Basil hatim and Jeremy Munday. Translation An advanced resource book..p :5
[8] European Association for Machine Translation. Proceedings of the 13th Annual

Conference of the EAMT, pages 81

[9] Naftali Kadmon. Toponymy The Lore,Laws and Language of geographical Names..p:149
[10] The Concise Oxford English Dictionary

[11] Op cite
[12] Tatiana Gornostay and Inguna Skadiņa English-Latvian Toponym Processing:Translation Strategies and Linguistic p: 81–87
[13] H. Douglas Brown. Principle of language learning and Teaching.
[14] Caroline Lavecchia. Les triggers inter-langues pour la Traduction Automatique Statistique. .p :37

[14] Opcite. p :157
[15] Ibid
[16] Thierry Grass La traduction comme appropriation : le cas des toponymes étrangers :667
[17]Melamed2000 MELAMED,I.D2000. Models of translational equivalencemong

words.Comput.Ling.26,2,221-249

Bibliography

- Naftali Kadmon. Toponymy The Lore,Laws and Language of geographical Names..Vantage press New York
- Meta : journal des traducteurs / Meta: Translators' Journal, vol. 51, n° 4, 2006, p. 660-670.
- Motivations for Naming :A Toponymic Typology.ANPS Technical paper No. 2 2009
- A standardized Geographic Feature Catalogue for toponymic research ANPS Technical Paper n 1 2008.Australian National Place names Survey .Linguistics Department Macquarie University Sydney NSW 2109© 2008, 2009
- Les triggers inter-langues pour la Traduction Automatique Statistique. Thèse

présentée et soutenue publiquement le 23 juin 2010pour l'obtention du Doctorat de l'université Nancy2 (spécialité informatique)par : Caroline Lavecchia

- Basil Hatim and Jeremy Munday Translation .An advanced resource book.

First published 2004 by Routledge. Park Square, Milton Park, Abingdon, Oxon14 4RNISBN 0-203-50188-8 Master e-book ISBN

- Tatiana Gornostay and Inguna Skadiņa.English-Latvian Toponym Processing:Translation Strategies and Linguistic Pattern2009 European Association for Machine Translation.Proceedings of the 13th Annual Conference of the EAMT,Barcelona, May 2009
- Thierry Grass. La traduction comme appropriation : le cas des toponymes étrangers . Volume 51, numéro 4, décembre 2006, p. 660-670 Direction de la rédaction : André Clas (directeur) Éditeur : Les Presses de l'Université de Montréal ISSN : 0026-0452 (imprimé) 1492-1421 (numérique).
- Jean Louis vaxelaire. Pistes pour une nouvelle approche de la traduction automatique des noms propres. Université Paris VII, Paris, France Meta : journal des traducteurs / Meta: Translators' Journal, vol. 51, n° 4, 2006, p. 719-738.
- Authoul Abdulhay Lidilem et Olivier Kraif Lidilem. Alignment Of Bilingual Anthroponyms And Toponyms OPONYMS In Parallel Corpora French Arabic. --Université Stendhal Grenoble3 Grenoble, France, 38000 Article:p2
- Melamed2000 MELAMED,I.D2000.Models of translational equivalence mong words.Comput.Ling.26,2,221-249
- H. Douglas Brown. Principle of language learning and Teaching.. Fourth edition. Copy right 2000 by Addison Wesley Longman ,Inc A Pearson Education Company ISBN0-13-017816-0

***

Annex

AF : A landing area for aircraft

ANCH: A sheltered location suitable for vessels to anchor.

BATH. Any feature of marine waters which is always or usually submerged [Temporary set;

ultimately to be replaced by a number of feature sets dominated by + MARINE,

+BATHYMETRIC

A concave recess made by the sea in the coastline, larger than <COVE> but

smaller and more convex than <BGHT>
BAY. Included terms: bay1
. Aero relates to air, as opposed to ground, transport,
Agricultural :relates to farms and other agricultural establishments
Abattoir: A building in which animals are slaughtered for their meat and byproducts.
Feature set: <BLDG>
Abbey (A building or buildings occupied by monks or nuns under an abbot or abbess. Feature set: <BLDG>)
aboriginal outstation(An area of land on which a small community of Aboriginal
People lives away from larger settlements. = out camp, outstation. Feature
set: <HMSD>) [1]
Aero: relates to air, as opposed to ground, transport
Agricultural: relates to farms and other agricultural establishments
Apical: is recognized as the uppermost part of a larger relief feature
Arboreal : is characterized by the presence of tree-like vegetation
Barrier,: serves to contain or to obstruct passage
Bathymetric: is a feature of marine or non-marine waters which is always or
predominantly submerged
