



**KAUNAS UNIVERSITY OF TECHNOLOGY**  
**FACULTY OF SOCIAL SCIENCES, ARTS AND HUMANITIES**

**Kęstutis Ivanauskas**

**CROWDSOURCING AS A TOOL FOR LOCALIZING WEBSITES**

Master's Thesis

**Supervisor**  
Prof. Vilmantę Liubiniene

**KAUNAS, 2017**

**KAUNO TECHNOLOGIJOS UNIVERSITETAS**  
**SOCIALINIŲ, HUMANITARINIŲ MOKSLŲ IR MENŲ FAKULTETAS**

**Kęstutis Ivanauskas**

**SUTELKTINIS VERTIMAS – KAIP INTERNETO SVETAINIŲ  
LOKALIZAVIMO PRIEMONĖ**

Baigiamasis magistro projektas  
**Technikos kalbos vertimas ir lokalizacija (kodas 621U60002)**

**Vadovas**  
Prof. Vilmantė Liubinienė

**Recenzentas**  
Doc. Ramunė Kasperavičienė

**Projektą atliko**  
Kęstutis Ivanauskas

**KAUNAS, 2017**

KAUNO TECHNOLOGIJOS UNIVERSITETAS  
SOCIALINIŲ, HUMANITARINIŲ MOKSLŲ IR MENŲ FAKULTETAS

Kęstutis Ivanauskas  
Technikos kalbos vertimas ir lokalizacija, 2 kursas

Baigiamojo projekto „Sutelktinis vertimas – kaip interneto svetainių lokalizavimo priemonė“

**AKADEMINIO SAŽININGUMO DEKLARACIJA**

2017-05-25

Kaunas

Patvirtinu, kad mano Kęstučio Ivanausko baigiamasis projektas tema „Sutelktinis vertimas – kaip interneto svetainių lokalizavimo priemonė“ yra parašytas visiškai savarankiškai, o visi pateikti duomenys ar tyrimų rezultatai yra teisingi ir gauti sąžiningai. Šiame darbe nei viena dalis nėra plagijuota nuo jokių spausdintinių ar internetinių šaltinių, visos kitų šaltinių tiesioginės ir netiesioginės citatos nurodytos literatūros nuorodose. Įstatymų nenumatytų piniginių sumų už šį darbą niekam nesu mokėjęs.

Aš suprantu, kad išaiškėjus nesąžiningumo faktui, man bus taikomos nuobaudos, remiantis Kauno technologijos universitete galiojančia tvarka.

---

(vardą ir pavardę įrašyti ranka)

## LIST OF FIGURES

Figure 1. The different components of web localization quality .....	19
Figure 2. Entire blocks of untranslated text .....	31
Figure 3. String ready for translation on <i>Facebook</i> localization page .....	33
Figure 4. Localized calendar on <i>Facebook</i> website .....	35
Figure 5. Localized date and time format on <i>Facebook</i> website .....	35
Figure 6. Localized calendar on <i>Google+</i> website .....	36
Figure 7. The distribution of errors in localization of <i>Facebook</i> .....	38

# CONTENTS

LIST OF FIGURES.....	4
CONTENTS.....	5
SANTRAUKA.....	6
SUMMARY.....	8
INTRODUCTION.....	10
1. PECULIARITIES AND STRATEGIES OF WEBSITE LOCALISATION.....	11
1.1. Digital genres.....	11
1.2. Layers of localisation.....	13
1.3. Aspects of website localization.....	15
1.4. Assessing the quality of localized websites.....	17
1.5. Phenomenon of crowdsourcing.....	21
1.5.1. The emergence of crowdsourcing.....	21
1.5.2. Ethical issues.....	23
2. PRACTICAL ANALYSIS OF <i>FACEBOOK</i> AND <i>GOOGLE+</i> .....	26
2.1. Methodology.....	26
2.2. Comparative analysis of <i>Facebook</i> and <i>Google+</i> .....	26
2.2.1 Textual errors.....	26
2.2.2 Linguistic errors.....	31
2.2.3 Pragmatic errors.....	37
2.3. External quality: web usability, functionality and compliance with client's commission.....	37
2.4. Results of investigation.....	38
CONCLUSIONS.....	40
REFERENCES.....	41

Ivanauskas, Kęstutis. Sutelktinis vertimas – kaip interneto svetainių lokalizavimo priemonė. *Magistro* baigiamasis projektas / vadovas prof. Vilmantė Liubinienė; Kauno technologijos universitetas, Socialinių, humanitarinių mokslų ir menų fakultetas.

Mokslo kryptis ir sritis: 04H Filologija.

Reikšminiai žodžiai: *vertimas, lokalizacija, svetainės, internetas, sutelktinis vertimas*

Kaunas, 2017. 42 psl.

## SANTRAUKA

Šiame darbe tiriama sutelktinio vertimo kaip naujos vertimo rūšies pritaikymas interneto svetainėms lokalizuoti. Sutelktinis darbas, kurį dažniausiai atlieka neprofesionalai, gerai išmanantys tam tikrą sritį ir nereikalaujantys už tai atlygio, pasitelkiamas įvairiose srityse – nuo operacinių sistemų kūrimo iki verslo idėjų pasiūlymų ir įvairių problemų sprendimo. Tokia darbo rūšis įmonėms leidžia sutaupyti nemažai lėšų, todėl susidomėjimas ja nuolat auga. Vertimo ir lokalizacijos srityje sutelktinis vertimas dažniau sutinkamas lokalizuojant žaidimus ar kaip mėgėjiško subtitrų vertimo atšaka, tačiau po truputį toks vertimo būdas pradamas taikyti ir verčiant bei lokalizuojant interneto svetaines. Bet čia pat svarbiu tampa ir darbo kokybės klausimas. Ar sutelktinio vertimo kokybė gali konkuruoti su profesionalių vertimų kokybe? Jei taip, ar vertimui, kaip profesijai, kyla kokių nors grėsmių?

Nors sutelktinio vertimo populiarumas pasaulyje paskutiniu metu auga, mokslinių tyrimų interneto svetainių vertimo srityje dar nėra itin daug, todėl juntamas jų poreikis. Projekto tikslas – iširti sutelktinio vertimo, kaip priemonės, panaudojimą interneto svetainėms lokalizuoti. Šiam tikslui pasiekti išsikelti tokie uždaviniai: nustatyti kriterijai pasirinktų interneto svetainių lokalizavimo kokybei įvertinti, įvertinta sutelktinio vertimo būdu lokalizuotos svetainės kokybė, sutelktinio vertimo būdu lokalizuota svetainė palyginta su profesionalaus vertimo biuro lokalizuota svetaine. Darbas sudarytas iš dviejų dalių – teorinės ir empirinės. Teorinėje dalyje aptarti skaitmeniniai žanrai, išskirti jų tipai, taip pat aprašyti lokalizacijos sluoksniai ir pateikti aspektai, į kuriuos būtina atsižvelgti lokalizuojant bet kokią internete patalpintą svetainę. Taip pat pateikta lokalizacijos kokybei įvertinti skirta schema, nurodyti esminiai akcentai, į kuriuos būtina atsižvelgti vertinant kokybę. Išsiaiškinta, kad interneto svetainių lokalizacijos kokybė skirstoma į išorinę ir vidinę. Vidinė kokybė apimanti tekstinius, lingvistinius ir pragmatinius aspektus, o išorinė – funkcionalumą, naudojimąsi internetu ir kliento užsakymo metu pateiktus nurodymus. Empirinėje dalyje palygintos dvi socialinių tinklų svetainės – *Facebook*, kuri lokalizuota sutelktinio vertimo būdu, ir *Google+*, kurią lokalizavo profesionalus vertimo biuras. Atliekant analizę didžiausias dėmesys skirtas *Facebook* socialinio tinklo svetainės vartotojo sąsajos elementų ir įvairių meniu vertimui, kurio kokybė palyginta su panašiais *Google+* svetainės elementų vertimais. Išanalizavus tirtus pavyzdžius sutelktinio vertimo būdu lokalizuotoje svetainėje rasta rimtų klaidų: nesuderinta terminija, neišlaikomas vienodas formalumo

lygis, neišversti dideli teksto vienetai arba vertimas yra tik dalinis. Taip pat pastebėta esminių linksnių vartojimo klaidų, klaidingo vertimo atvejų, kai reikšmė stipriai iškraipoma, netinkamai lokalizuota data ir, iš dalies, kalendorius, kitų lietuvių kalbos normas neatitinkančių klaidų. Tuo tarpu profesionaliai lokalizuotoje svetainėje didesnių klaidų nepastebėta, aiškiai juntamas skirtumas tarp šių dviejų interneto svetainių lokalizavimo kokybės. Atsižvelgus į gautus rezultatus tyrimo pradžioje išsikelta hipotezė, kad lokalizacijos, atliktos sutelktinio vertimo būdu, kokybė yra prastesnė už profesionalaus vertimo biuro pateiktą kokybę, patvirtina. Be to, pastebėta, kad objektyviai įvertinti išorinę kokybę dar sudėtinga, nes tai gana naujas dalykas ir nėra šaltinių, kuriais remiantis būtų galima tokį vertinimą atlikti.

Ivanauskas, Kęstutis. *Crowdsourcing as a Tool for Localizing Websites: Master's thesis* in Translation and Localization of Technical Texts / supervisor prof. Vilmantė Liubiniėnė. The Faculty of Social Sciences, Arts and Humanities, Kaunas University of Technology.

Research area and field: 04H Philology

Key words: *translation, localization, web, websites, internet, crowdsourcing*  
Kaunas, 2017. 42 pages.

## SUMMARY

The popularity of crowdsourcing has been growing more and more in recent years. It is special type of phenomenon when people are working for their own sake and are not willing to get any type of award as their goal is being part of community and contributing to various types of tasks. Crowdsourcing is met in different spheres, starting from creation of operational systems and finishing with proposing ideas for business or solving business-related problems. Such type of work let companies save a lot of money and use them in other fields. In terms of translation of localization, this phenomenon is met mostly in game localization or as a type of translating films where all the work is being done by fans. The latter is called *fansubbing*. Recently, crowdsourcing is more widely used for translating or localizing websites as well. Despite that work done by the crowd raises questions on the quality of a final product. Is it the same in comparison to professional translation? If so, does this poses any threats to translation as profession?

Even though the popularity of crowdsourcing is seen to be growing in recent years, in terms of web localization there are not so many researches done and thus the need of such investigation is greatly felt. The aim of this work is to research crowdsourcing as a potential tool for localization of websites. To achieve this aim the following objectives have been set out: to set out criteria for evaluation of the selected websites localization quality; to evaluate the quality of the website localized by means of crowdsourcing; to compare the quality of translation in the website localized by employing crowdsourcing with the one translated by professional translators. The paper consists of two parts – theoretical and practical. In the theoretical part such subjects as digital genres and their types has been discussed, also layers of localization and the most important aspects of web localization that should be taken into consideration presented. Moreover, the model for quality evaluation has been provided. It has been clarified that quality in web localization is distinguished into external and internal one. Internal comprises textual, linguistic and pragmatic factors, while internal deals with functionality web usability and compliance with client's commission. In the practical parts two social networking sites, such as *Facebook* and *Google+* have been analysed and compared. The former site has been translated by means of crowdsourcing and for localization of the latter one professional translation agency was responsible. During the analysis all the focus was devoted to translation of user interface elements and menus on *Facebook* website and their quality compared with those on *Google+*



website. Having completed the analysis of a crowdsourced website some serious errors in translation has been found. The localized version of *Facebook* contained lots of inconsistencies, different level of formality, large pieces of untranslated or partially translated text. Moreover, some serious case usage errors as well as mistranslations, crude shifts in meaning and other impermissible errors have been found. Meanwhile, the site localized by translation agency did not contain such errors and the difference in quality between two sites was greatly felt. The hypothesis raised in the beginning of work that the quality of websites, localized by crowdsourcers, is not as high as of those localized by professional translators has been approved. Finally, it has been noticed that the assessment of an external quality can be done only on a very subjective level, since it is still a relatively new thing and there are simply no sources which would help during such evaluation.

## INTRODUCTION

In the modern world companies are made to go worldwide in order to be successful. This in one way could be achieved with a help of Internet websites advertising companies' services or products. But this is not that simple as it may look like, since there are many different countries and many different languages in the world which makes communication or the transfer of a message encumbered. In order to promote products or services via the web one has somehow to overcome all those obstacles. One of ways to do that is translation. However, when it comes to websites, translation is not enough. In order to communicate a specific message it has to be adapted to people of that specific culture or, in other words, it has to be localized. But in web localization the message can be transferred not only by the means of language – some specific peculiarities should be considered as well. As localization is not a cheap process, companies tend to look for other economically useful ways of making their websites known to the world. One of ways to do that is by employing crowdsourcing – quickly growing phenomenon of making the crowd voluntarily work for you and reach the intended purpose.

The aim of this paper is to research crowdsourcing as a potential tool for localization of websites. To achieve the above indicated aim the following objectives have been set out:

- To set out criteria for evaluation of the selected websites localization quality;
- To evaluate the quality of the website localized by means of crowdsourcing;
- To compare the quality of translation in the website localized by employing crowdsourcing with the one translated by professional translators.
- Research object – English and Lithuanian versions of social websites of *Facebook* and *Google+*.

Research hypothesis – the quality of websites, localized by crowdsourcers, is not as high as of those localized by professional translators.

Chapter One “Peculiarities and Strategies of Website Localization” presents strategies which every translator/localizer must take into consideration while translating websites. Also, understanding of digital genres, aspects of localization in general, assessment of quality is given and crowdsourcing is presented. Chapter Two “Practical Analysis of *Facebook* and *Google+*” encompasses results of translation quality found in *Facebook* and *Google+*. Moreover, examples are analysed and evaluation of quality is given on textual, linguistic and pragmatic levels. Finally, the MA thesis is supported by conclusions.

# 1. PECULIARITIES AND STRATEGIES OF WEBSITE LOCALISATION

In the upcoming sub-chapters various website localization properties are going to be presented. All the focus is going to be devoted on genres in the web localization and localizable elements as well as a scheme for quality evaluation is going to be proposed. Last but not least crowdsourcing and ethical issues that it poses is going to be presented.

## 1.1. Digital genres

From the very first day of web existence, it easily adopted textual genres used in newspapers as well as offered new possibilities of adapting them and gave way for such sites as social networking sites, wikisites or videoblogs to emerge. This fact puts it very clearly how a web genre is created. Usually a very well-known textual genre is being improved to a more sophisticated level, more suitable to the needs of humanity of the modern times. These web genres or cybergenres now are generally called digital genres, i. e. covering all the genres possible on the Internet. Introducing this term Jimenez-Crespo in his book *Translation and Web Localization* provides the definition suggested by Erickson who defines digital genre as “A patterning of communication created by a combination of the individual (cognitive), social, and technical implicit in a recurring communicative situation” (Erickson qtd. in Jimenez-Crespo, 2013, p. 79). One of the most peculiar features of these genres is their rapid emergence over the other ones, as web functionality is improving almost every day. In general all genres are improving constantly, however improvement of the digital ones is much more difficult to be controlled and foreseen. Moreover, these genres may also disappear just like any other and in some cases, as Jimenez-Crespo states, digital genres “life-expectancy might be dramatically shortened” (Jimenez-Crespo, 2013, p. 78).

Social networking site and the professional/academic homepage are the two distinctive genres that the digital one quickly evolved into. In case of the former, it encompasses different goals, such as professional, romantic and others, hence popularity of social networking sites is enormously growing. Even though academic personal pages still remain quite a popular digital genre. In fact, there are two main types of digital genres – *extant* and *novel*. Extant are described as those that were transferred to the Internet without any adaptations or as they appear in the source media, while novel fully depend on their new medium, whether it be search engines or videoblogs (ibid, p. 79). Further on extant genres are subdivided into *replicated* and *variant*. Replicated are those “made available online without any adaptations” and since they are hardly adaptable to the medium, they are not successful in terms of the web unlike variants that “show some minor adaptations” (ibid). What regards novel digital genres is that they are subdivided into *emergent* and *spontaneous*. Emergent genre evolves from a printed one,

for instance online encyclopaedia, and features various functionalities, while spontaneous are the ones that appears right on the new medium, for example as blogs or homepages (ibid).

Whatever the genre it is, it encompasses a definite communicative situation determined by audience, field of usage, mode and direction. Viewers of any company website expect a formal tone to be met in such site, while their alternative social networking sites would include more informal one due to particular communicative purposes to be achieved among different users. On the Internet, in general, prevails two forms of interpersonal communication – synchronous and asynchronous. The former would be chats, videoconferences and similar types of communication while the latter covers emails, postings and so on. However, in some social networking websites, for example *Facebook*, such conversation by using chats and emails have been merged into one platform and have peculiarities of both spoken language (chats) and written forms of communication (emails) (ibid). Some scholars, like Ana Janoschka, states that websites now present new model of communication known as *interactive communication process* where interactivity lies between users, senders and medium (Janoschka, 2003). This process is characterized as meeting the same principles and criteria of mass communication found in advertising or television, hence it is also a mass communication process. Moreover, in web localization interactions happens not only between humans, but between humans and computers (or other devices) as well. These two types of processes are known as *interaction* and *interactivity* (ibid). Interaction is such a communicative process that happens between two human beings, i.e. among the sender and the receiver while reading websites, emails, chats, various forms and so on. Interactivity is a bit different process, as it happens not between two humans, but rather between a human and a medium when users utilize search functions, activate hyperlinks, uses navigation menus, etc. Having taken all these factors into consideration it is possible to state that website localization is asynchronous mass-communication process contributing to communication among users and websites. However, it has been noticed recently, that this process also facilitates communication between users themselves as they translate content generated by others. Communication itself is directed from a company or organization to audience, though websites provide interactivity options, such as forms, subscriptions to newsletters and so on (Jimenez-Crespo, 2013). Audience is also divided into primary, supervisory and peripheral (Jeney, 2007). In terms of websites, their goal is to reach the primary audience, while supervisory refers to commissioners or translation agency responsible for localization. The last type of audiences denotes visitors who visit the site, even though they have not been specifically attracted by it (ibid).

Finally, genres on the web are divided into *informational*, *communicative or interactive*, *instrumental*, *advertising* and *entertainment genres* according to the purpose a website is meant to achieve: to provide information, to assist in communication or interaction, to use it as a tool or instrument, to advertise products or services, to provide entertainment (Jimenez-Crespo, 2013).

Definitely, websites may encompass several enlisted goals. In case of social networking sites, these belong to communication-interaction genre and carry two primary functions – expositive and exhortative. Moreover, they cover general, professional and personal interests’ subgenres (ibid).

## 1.2. Layers of localisation

Gerhard Chroust in his article *Software like a Courteous Butler-Issues of Localisation under Cultural Diversity* put emphasis not on the web localization in particular, but described some specific features of software localization, where culture plays one of the most important roles in the process of localization. Although software presents different type of media, some approaches to software localization can also be adapted to translation of websites. First of all Chroust gives the definition of localization, borrowing it from the GSSI, which defines it as “the process of adapting a product to reflect the local standards, culture and language of another market or the infusion of a specific culture into an international product” (GSSI qtd. in Chroust, 2007, p. 3). However, websites may also reflect some cultural aspects typical to one or another part of the world. Chroust calls these aspects as layers and enlists seven layers of localization (Chroust, 2007):

- a) Cultural Layer;
- b) Social and Communication Layer;
- c) Business Conventions and Practices Layer;
- d) Graphic and Iconic Representation Layer;
- e) Semantic Layer;
- f) Grammatical Layer;
- g) Technological Infrastructure.

The basis of *Technological Infrastructure Layer* is technical and organizational provisions. Attention is devoted to the separation of text and code, keeping enough storage space for texts, proper coding of characters, using correct sort order, dealing with two-byte languages used in Asia and guaranteeing that reading and writing direction is correct (left-to-right, right to left). Moreover, adaptation of date, currency, time, word order of that specific locale is also being taken into consideration. Meanwhile *Grammatical Layer* is meant for ensuring that technical languages conform to the required rules, texts are homogenous and standard, since many sentences are generated by a computer. Another type of layer called *Semantic Layer* deals, as the name suggests, with semantic peculiarities of text, i.e. with expressiveness of languages, abbreviations etc. This layer relies more on translations of the human language. Above the Semantic Layer goes *Graphic and Iconic Representation Layer* which primary concern is the usage of correct symbols, colours, taboos, body language, private setting and so on (cf. ibid, 2007).

The rest three layers are basically interrelated with cultural attributes. Talking about *Business Conventions and Practices Layer* Chroust states that leadership approach differs in various cultures and within organizations according to the type of organization (democratic, authoritarian, participative) and that “management support tools have to conform in their users’ expectations in their interfaces and procedures” (Chroust, 2007, p. 7). The scholar also implies that the structure of interfaces, websites and pictorial representation must be adapted to cultural expectations depending on type of the society (egalitarian or hierarchical). This should be also reflected in the design of web pages. Another aspect which should be taken into account is dates and deadlines that varies according to culture. To make project management tools acceptable, deadlines have to be clearly stated. Furthermore, performance may also be understood in many different ways depending on whether the country is individualistic or collectivistic and inner or outer directed.

By *Transactions Layer* various social and communicational aspects are meant. One should be very attentive and choose the right addressing and greeting style as well as gestures or the correct level of familiarity, since all of these features may be reflected by pictures and animation. Another important aspect is the way people agrees or disagrees. A localizer localising any material for Asian countries should bear in mind that some of these nations avoid the direct disagreement expressed by ‘no’ and sometimes ‘yes’ may mean disagreement too. Chroust then carries on talking about communication styles. He puts emphasis on high and low-context cultures, stating that for high-context cultures, such as Asian or Hispanic, the context and non-verbal aspects are taken into account, while the other group of culture expects explicit communication and some types of verbal interaction should be used as well. The author also pays attention to social classes and claims that while localizing any localizer must consider such aspects as age, education and other peculiarities that could be important for individuals of some subculture or class. In this case the message should be adapted according to different social stratum or hierarchy, if the society is highly hierarchical. What is more, gender should not be forgotten, since the role of women varies in different countries. In Western countries it would be no surprise to see a woman taking the highest positions at any occupation, while in Arabic culture this would not be accepted at all. Therefore in some national languages the distribution between gender-dependent and gender-independent words is of high importance and must be very well considered.

The last, but not least *Cultural Level*, as its name suggests, deals with aspects embedded in the people’s culture. In this case a localizer has to bear in mind such culture-dependent attributes as taboos, puns, metaphors, jargons and humour. These peculiarities may play a crucial role on accepting various products. But one should also be very attentive, since, for example, taboos are highly sensitive and the usage of these words may result in a serious negative effect. Chroust provides an example when some Danish caricatures of the prophet Mohammed in the media caused riots in Islamic

countries because residents of these nations interpreted it as an insult, which in this part of the world is taken very seriously (ibid). In case of metaphors, puns and jargons situation is actually the same, only here translator or localizer must use such expressions that the nation is familiar with and avoid those who may result in misunderstandings and burden the communication. Humour is especially difficult to be conveyed. It is so deeply interrelated with various dominant contexts in one or another country that very often it is not even translated and in most cases transcreation is the only way out. In this case some specific type of humour that entails the same or similar effect should be taken from the context of that definite region.

Chroust also provides some possible results of poor localization. According to him, poor localization may result in reduced effectiveness and productivity or may cause anger and offences due to misinterpretation of signs or environment (cf. ibid, 2007). In the worst cases this can lead to broken personal or economic ties or even rejection of products. Therefore it is suggested to spend more on localization and have a satisfactory result, which would help to enter the global market successfully.

### 1.3. Aspects of website localization

There are many important factors that should be taken into account when localizing a website. The model provided by Chroust served as a starting point for other researchers who took a deeper look into web localization. One of those was Anthony Pym. In his article *Website Localisation* he introduces his own approach to this process. First of all, the author points out the process of internationalization, which is a process of making the website to be neutral, functional and constructed in a particular way that users in different cultural locales could find it attractive (cf. Pym, 2010b, p. 5). In terms of localization, the emphasis is not only on textual elements – non-textual ones must be localized too. Pym provides a list of constituents that comprises content that are to be localized and some of them are: songs and music, address formats, iconic conventions, colours, name and date formats and so on.

Localization of websites very much depends on their nature, since sites can be monolingual, bilingual or multilingual or, in other words, there can be one, two or several languages available, but the format may be retained. It depends on which marketing strategy the company is going to apply. Pym enlists five degrees of localization, identified by Singh & Pereira, which are as follows: standardized, semi-localized, localized, highly-localized and culturally customized (ibid). ‘Standardized’ means that the website in all the countries has a unique format and is not translated, while ‘semi-localized’ defines the site that contains information about many countries. The degree ‘localized’ should be given to the site which is fully translated in each country, ‘highly localised’ – to the one that is translated and culturally adapted and ‘culturally customized’ – to absolutely new site, containing all the specific attributes of that culture. According to the research done by Singh & Pereira

in 2005, most of the websites were given the degree of ‘localized’ and only 17 percent of them were ‘standardized’ (Pym, 2010b). Sometimes websites are translated, but some specific sort of technical information is still not conveyed, since it is intended for highly specialized users.

Another very important aspect that should be taken into account deals with the way users read websites. According to the research made by Nielsen it is claimed that reading from a computer screen is about 25 percent slower than reading from a paper. This phenomenon has been explained by the fact that websites are a “non-linear means of communication” (ibid) and users look through the website very quickly only casting their eyes on some particular items of their interests. Moreover, as Nielsen’s studies show, people tend to read web pages starting “at the top-left and look across the page horizontally in one or two sweeps, and then skim down the page vertically, giving an F-shape pattern” (ibid). It is also presumed that even 17 percent of users do not read one page longer than ten seconds and even educated persons read just 20-28 percent of information provided on that particular page. Therefore it becomes clear that pages on the web cannot be written in the same manner as standard, printed editions of media. Web pages are to be used, not simply being read. From this perspective some parts that are not going to be looked through attentively can even be translated using machine translation, while the most readable spots should be conveyed with great attention. Moreover, Pym suggests that web-based machine translations help website users to catch the idea of what that page offers or what it is about (Pym, 2010b, p. 9). Of course, in this case quality is not the primary goal. It is the cheapest way of translation.

Finally, the design of a website can even be more important than the text itself, since, as it has already been mentioned, people do not tend to read web pages. Even if translations are of very high quality, poor design may not attract users and vice versa. The function of a website should always be considered in the first place. If it is a site of the company selling clothes or shoes or similar things, naturally there are going to be less text and more pictures and usually these graphical elements are going to be big and detailed, while the text can be small and even hard to read. On such a site text may not be written or translated well, unless it is a slogan. On the contrary, news sites are going to have much text and less or smaller pictures, therefore it is expected that the quality of texts is going to be very high and letters should be big enough to read. Again, what kind of design is going to be used depends on locale and its cultural expectations. One should remember that a nice website with all the graphical elements and with a sufficient number of visitors in one part of the world may not be so successful in the other. However, everything depends on money that a company wants to spend for localization. The first step before the start of localization is to examine the locale and see if it is worth investing here. Adapting a website to its cultural and other needs may require quite a big amount of money, thus the goal and function of the website should be very clearly stated. Maybe the machine translation with some major editing would be enough.



#### 1.4. Assessing the quality of localized websites

Web localization is a composite process, involving many different specialists, such as engineers, managers, terminologists, quality-assurance operators and others. In addition to this, it is a process, which can be carried out by large corporations as well as by medium or small ones or even one individual can perform all the necessary work. However, all the efforts put into adapting the website for a specific target culture must somehow be evaluated. At first sight this may seem as a pretty simple task, but when ones tries to do that some obstacles appear. First of all, translation agencies, freelance translators or translation technology market perceives quality in different ways. Therefore how quality of the website is going to be evaluated should be decided before the start of localization. Nevertheless, while assessing the quality culture should be the primary and the main concern above all. Culture is rather a broad term, but in terms of web localization it covers several factors. Gibb & Matthaiakis in their article *A framework for assessing web site localization* enlists 13 of them in total. They all are of particular importance and should not be ignored at any cost. These are as follows (Gibb & Matthaiakis, 2007):

- Date order: one should take particular attention to the localized format of date during quality assessment, as various countries uses different formats, such as DD/MM/YY in the UK, MM/DD/YY in the USA or YY/MM/DD in other regions. Time conventions should be considered as well, since in some countries it is agreed to utilize 24-hour format, while the US or UK resident would expect to see time provided in 12-hour clock with AM or PM beside the digits. Misleading formats, first of all, may cause serious problems for any person or company and, secondly, may lower such site’s usability by a huge margin or the site may not be used at all.

- Currency: all prices must be converted to the currency used in that locale according to the relevant exchange rate. However, Gibb & Matthaiakis highlights that this may cause a side effect, as physiologically appealing prices, for instance 9.99 instead of 10.00, after conversion may become unattractive to the customers, therefore it is up to that client to decide how to deal with prices in the best possible way.

- Measures: metric and imperial measures need to be converted, also paper sizes, clothes sizes and presentation of numbers as well.

- Characters: numbers denoting centuries may need to be changed. For instance, 17 century may be written as XVII, but this convention is more or less nation-depended, therefore particular attention should be taken as to which format is required. In addition to that, punctuation should also be taken into consideration, as in Greek, for example, “;” means the same as “?”.

- Examples: when dealing with examples one has to provide easily denotable ones for that

specific locale and context. For example, *Renault* in France, *BMW* in Germany, *Inter Milan* in Italy, *Barcelona* in Spain. Moreover, even traffic lights might need to be changed, as in the USA the scheme is red, yellow and green, while in Japan it is red, yellow and blue.

- Language: all the foreign language text must be translated and the usage of idiomatic expressions should be lowered to minimum as this would allow the content to be translated much quicker. Furthermore, source text writers must bear in mind that a translated text is usually longer, therefore it must be of respectful length as some more space may be needed in another language and in case of mobile devices it may not fit properly on the screen. If the website contains any video content, it needs to be dubbed or subtitled.

- Colours: it must be checked if their usage match the appropriate locale, since the same colour may have different meanings in various regions, e. g. red for Chinese stands for joy, but in any other western country it would be seen as a danger.

- Icons: it should be carefully reviewed, if they pertain to that specific culture and can be easily denoted. Such icons as mailboxes, waste bins, house styles and so on differ in the very same manner as colours do.

- Layout: this factor is very important when a website has to be localized for a rather specific locale where readers tend to read pages in a different way, for instance, from right to left and then from top to bottom as this would be in case of Hebrew.

- Images: they can contain different types of content which in some cases can be very sensitive, therefore they must be carefully reviewed and their familiarity checked.

- Legal requirements: this factor is very important as almost every country apply different requirements (or in a slightly different way), especially those related with privacy. In some countries, for instance, in the Netherlands, any direct comparisons with competitive products are allowed.

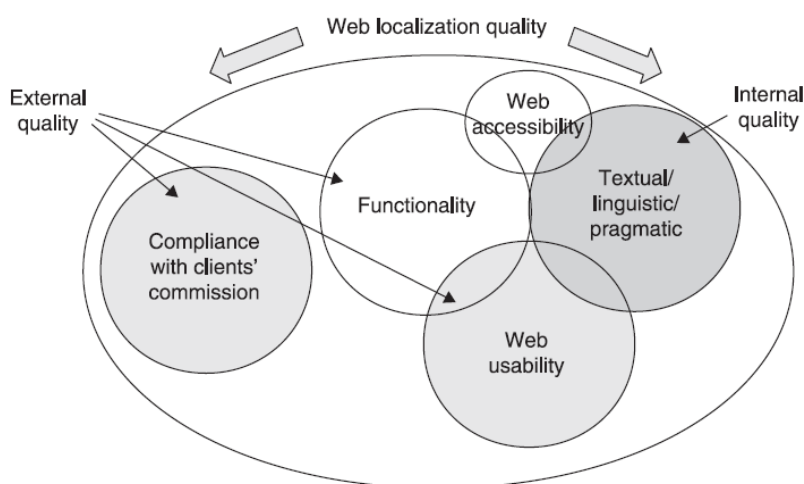
- Age group: language must be adapted for to the relevant audience's needs.

- Experience: content might be chosen as to match different levels of expertise. It should also not to be forgotten that certain cultures have different experience of perceiving things, just like it is in case with Eskimos perception of snow that can be designated in far more numerous way anyone could imagine.

However, one of the biggest issues related with quality is that web localization itself is “a relatively new phenomenon [...] and still has no set of canonized criteria of its evaluation” (Wright qtd. in Jimenez-Crespo, 2013, p. 103). Thus, some methods are to be presented and implemented. One of the ways to do that is using quality assurance (QA) or quality control (QC) strategies. The latter is used to maintain quality during all the stages of the process from the beginning till the very end when the final product is checked. What is more, quality is distinguished between external and internal. Internal quality shows how well all those intrinsic issues within text have been solved, while external

describes how well the intended purpose has been reached and whether it satisfies the implied client (cf., *ibid*, p. 105). Moreover, external quality is grounded on various standards that each country has. For example, in Europe European translation standard EN 15038 is used, while in North America – ASTM 2575-06. Also, there is an international standard called ISO 9000. Nevertheless, Jimenez-Crespo claims that “quality evaluation is bound and constrained process” (*ibid*, p. 108), encompassing two types of constraints: contextual and procedural. Budget, time for completion, cultural norms of quality – these are just few of contextual ones, while procedural constraints depends mostly on QA system used, translation-localization technology applied, cognitive constraints and others. In an ideal case, the quality might be seen as a balance between contextual and procedural constraints.

Taken into account all these factors, Spanish scholar provides a graphical representation of factors on which the quality of website should be grounded. He introduces a rather general illustration of the most prominent aspects on which web localization quality is grounded:



**Figure 1.** The different components of web localization quality (Jimenez-Crespo, 2013)

As it is seen from the model, the quality of web localization basically depends on two factors – external and internal. External quality is related with functionality, web usability and compliance with clients’ commission, while internal quality encompasses textual, linguistic and pragmatic qualities. All these fields, except compliance with clients’ commission, are more or less interdependent, which means that localized websites with some functional problems may reduce web usability and accessibility, despite the fact that linguistic qualities would be retained perfectly. On the other hand, low quality of textual/linguistic/pragmatic aspects may also result in reduced web usability and accessibility, even though functionality would not be affected so much, since the latter merely depends on various technical issues rather than linguistic ones. This distinction in one or another way also encompasses Chroust’s layers of localization. First of all, all those layers belong to internal quality of web localization. Then it can be assumed that Technological Infrastructure Layer belongs to linguistic and textual groups of aspects, while Semantic Layer would clearly go under the group of linguistic

peculiarities. Grammatical Layer would also cover textual and linguistic features as it deals with confirming to the language rules, homogeneity and standardization of texts. Nonetheless, Graphic and Iconic Representation Layer would likely go under the pragmatic level as well as Business Conventions and Practices Layer together with Social and Communication Layer. Last but not least Cultural Layer encompasses all the textual, linguistic and pragmatic aspects as they are more or less culture-dependant.

Pragmatic aspects in this case encompass politeness and various speech acts met in the process of communication between a user and a website. The feature of politeness is a culture-dependent element, the same like a custom to take or not to take the shoes off after entering into other people's house. In some countries staying with the shoes on is strictly unacceptable, while in others taking them off can be considered as an impolite or strange action (Archer, Aijmer & Wichmann, 2012). Thus, the degree of politeness or impoliteness depends on how that culture treats directness, i.e. expressing ideas directly to the speaker. In English, for example, different indirect requests may be used as not to express impoliteness while communicating with people of higher social status. However, the expression of politeness varies by language which is an integral part of a culture. In some cultures it can be indicated by adding a title to one's name or surname, for instance Mr/Mrs/Ms/Dr/Professor Smith (Archer et al., 2012). English contains only the pronoun 'you' which is used both in formal and informal situations, therefore English speaking communities tend to add the aforementioned titles. In other linguistic communities these titles are not so often used as the degree of politeness is conveyed using pronouns. Scholars call such usage as a "T/V (tu/vous) distinction" (ibid). The so-called T-form is used when addressing friends or family and the V-form is met in communication with strangers and is perceived as a sign of respect. To strengthen the formality or respect, pronominal forms can also be written from the capital letter, for example 'Sie' in German or 'Jūs' in Lithuanian. Which pronoun is to be used depends on the parameters of relative power and degree of solidarity (ibid). Nonetheless, sometimes it can be problematic to decide which form is more suitable. Then the decision is taken based on particular individual or common desires. Linguistic communities that are accustomed to the usage of a T/V distinction may feel awkward if only one pronoun, such as 'you', is used or even take such usage as being too familiar (ibid).

The expression of directness or indirectness occurs by using *direct* or *indirect* speech acts. When one says *Close the window!* it is perceived as a direct speech act, while the indirect act can be made by using modal verbs, for instance *Would you be so kind and close the window?* Websites in the most cases contain direct speech acts, though translation may mitigate directness depending on the culture as the site can be used by different users and such mitigation serves like a face-saving strategy, e.g. *Log in/Prisijunkite* (please login). The term 'face' encompasses such aspects as being honest, well behaved and being a member of "valued social groups and institutions in the community" (LoCastro,

2012). As a result every human-being hopes others will recognize their face as well as their needs by means of verbal communication and try not to present a threat to it.

### 1.5. Phenomenon of crowdsourcing

In the following subchapter the history of crowdsourcing is going to be presented and the broader view on the definition itself is going to be given. Despite the fact crowdsourcing can be employed in translation, it may present some ethical threats in terms of translators and this profession as well.

#### 1.5.1. The emergence of crowdsourcing

The term *crowdsourcing*, also known as *fan translation*, *volunteer translation* or *user-generated translation*, was introduced by Jeff Howe as a portmanteau of *crowd* and *outsourcing* in 2006 in the *Wired* magazine article *The Rise of Crowdsourcing*. Howe used this term to define a new phenomenon which, as he thought, would revolutionise a wide range of fields, and primarily the normal way of doing business. According to Howe, technological advances have made cheap consumer electronics an everyday reality, therefore the gap between professionals and amateurs has diminished, allowing companies to use talents of crowd. The IT revolution has contributed in utilizing the best of people skills regardless of their location (*Crowdsourcing Translation*, 2012). In addition to this, one of the main preconditions for crowdsourcing to spread is relatively cheap and easily accessed tools that help to send information just in a matter of seconds (ibid).

Crowdsourcing strategies are applied both outside and in the business world in various areas, such as software development, humanitarian aid or the audiovisual sector. Originally, crowdsourcing was defined as “the act of taking a job traditionally performed by a designated agent (usually an employee) and outsourcing it to an undefined, generally large group of people in the form of an open call” (Aris, H. & Din, M. M., 2016). As the definition states, the job is usually done by a large group of people, which presumes that they may not be professionals, but rather amateurs. What people working in a large group really seek for is an enjoyment, the feeling of being part of the community. This allows them, for instance, to come up with new various problem-solving methods and production designs. As James Surowiecki states, “the many are smarter than the few” and “groups are often smarter than the smartest people in them”. Since the communities are usually not for their efforts, they are completely free to choose their topics and tasks. Even economic trends cannot control this as the crowd is seeking to be recognised and respected, rather than to be given an adequate sum of money. Solutions taken by non-experts in the specific field present the biggest and the most remarkable difference between crowdsourcing and traditional problem-solving methods and it is also one of the main reasons why large corporations choose crowdsourcing even when they have highly qualified

experts in-house. With the help of this method, they can reach various groups of people possibly aiming to take a new perspective in solving different kinds of problems and, at the same time, introduce innovative solutions.

Even though this may seem as a new trend in the business world, the work of amateurs has always been essential for the advancement of science and culture. Right up to the beginning of the 19<sup>th</sup> century, most of the science-related work was exclusively carried out by enthusiastic amateurs most of whom being of aristocratic origin, therefore they could spend all the time and use any tools they needed in terms of intellectual activities without being given any monetary reward as this was not their primary concern. When it comes to the modern times, it becomes clear that the involvement of amateurs has been greatly boosted by the Internet and social media which turned communication into quick and almost a borderless process. With the help of IT people come into groups more easily, spread their message, find supporters and organize events or seeks for an international support. One of the best examples of how fruitful the community of people sharing and supporting the same idea can be is the birth of *Linux*. The operating system that initially was an one-man's idea have soon become a huge project, where all issues have been resolved and improvements made merely by the users of the system whenever someone suggests any change. Even though the system was introduced quite a long time ago, it still has been used by a number of people and has gained the role of a serious competitor even to such a company as *Microsoft* (*Crowdsourcing Translation*, 2012). This fact only proves how efficient the group of people guided by the same goal – open and free to anyone operating system – can be.

Businesses have also been adopting crowdsourcing more and more. Usually they choose to utilize this method whenever they need opinion of their customers. For instance, when they want the crowd to propose ideas for new products, features, clips, slogans or packages. One of such cases was when *Pepsi* decided to introduce new cans and asked people to submit their ideas on how those cans should look like. Other companies, like *Danone* or *Nespresso*, created platforms where users had a possibility to vote for a new flavour or choose from different advertising campaigns. *Dell* also introduced a platform, called *IdeaStorm*, simply to know what their customers actually need, so that they could reduce the number of intermediaries which later resulted in reduced prices as well (*Crowdsourcing Translation*, 2012). However, crowdsourcing can be utilised not only for business needs, but even for scientific purposes. This is what the company *InnoCentive* strives for. It helps other organizations in finding a solution for problems they cannot solve themselves by introducing scientific challenges (*Crowdsourcing Translation*, 2012). The company has millions of problem solvers who carry their work using a specific cloud-type platform. *InnoCentive* has been doing so well that they attracted attention of such companies like *NASA*, *Procter & Gamble*, *The Economist* and others. What make other companies choose them are profit, quick results and a chance to present a problem they

could not solve themselves to somebody else, who are keen to contribute in reaching for project's success.

In terms of content creation, one of the most prominent case is *Wikipedia*. This well-known encyclopaedia is written merely by volunteers and any anyone willing can create or edit or translate articles on any subject and in any language. To this day, as the number on the website's homepage claims, there are 5,402,775 articles written in English, not even mentioning the numbers in other languages. Since this is a product that has been created by anyone from all over the world, some has doubts on quality and credibility of those articles and tend to consider *Wikipedia* as unreliable source of information. However, according to results of the study completed in 2005, this encyclopedia had been acknowledged as reliable as *Britannica*<sup>1</sup>. Another quite popular website, *Duolingo*, is for those who want to learn a new language. The idea of this site is that people learn new language through translating texts that are gathered by crowdsourcers. At first users learn some words, then they have to use those words in the sentence and translate them. When they get competent enough, they start translating actual texts which the site monetizes. In terms of social media, like *Facebook* or *Twitter*, users are allowed to propose translations themselves. If they are unhappy about any parts, translations can be declined and edited according to the native speakers' needs. Some languages even have style guides, but it is still only up to users to use it or not. When *Facebook* gave way to this opportunity, users in some parts of the world responded so enthusiastically that the whole site translation was completed in two-weeks' time (Spanish case) or even in 24 hours (French case) (*Crowdsourcing Translation*, 2012).

Finally, the best example showing how people involved can be is the so-called *fansubbing*, i.e. subtitles that are created by the fans of movies. In some regions fan communities are so involved into the whole process that they even provide requirements for contributors to fulfil, have editors and reviewers – in other words, the whole job structure that one would find in a normal company. Moreover, some content has been translated so well, that it even contributed to success of that piece when it was actually broadcasted on TV (*Crowdsourcing Translation*, 2012).

### 1.5.2. Ethical issues

Although crowdsourcing seems to be a great way for problem-solving and decision-making processes in translation, this type of work also raise some ethical questions to translation as profession. Crowdsourcing as a phenomenon contributes to visibility of translation, shows its value to society and help minor languages become more visible online, but at the same time it poses a threat that the work involved in the translation process may become devalued and translators will not be treated as

---

<sup>1</sup> <https://www.cnet.com/news/study-wikipedia-as-accurate-as-britannica/>

professionals. Julie McDonough Dolmaya in her article *The Ethics of Crowdsourcing* enumerates three biggest issues, related with remuneration, translation visibility and the case with minor languages.

In the case of remuneration the biggest issue is that some people may not consider translation a true profession “because it does not seem to have a monopoly on a value goal that is not shared by other groups (unlike medicine, with the goal of health or law, with the goal of justice)” (McDonough Dolmaya, 2011). Since translation requires specific skills, then one should reasonably assume that translators deserve to be paid professional rates. For this reason, some professional translator associations, such as the *Asociación Guatemalteca de Intérpretes y Traductores* or the *Irish Translators and Interpreters Association*, oblige members to charge significantly higher fees than the specified ones to avoid accepting work at rates of pay unreasonably below market norms (ibid). Other similar associations advise members to charge “fair and reasonable” fees according to their experience, the difficulty of the text, and the importance of the project. However, it is indicated that all of these cannot prevent amateurs or non-members from setting themselves up as translators or how they should choose to be remunerated for translation work. But then the question arises is it ethical for an organization to look for volunteers to translate some content and not to offer any monetary reward? It is assumed that members can only accept below-market rates when translating for non-profit initiatives. But in these cases members are obliged to inform clients of the market value of the translation work. Moreover, various professional translator networks, which provide heavily discounted services to not-for-profit organizations, seek for translators “to volunteer their skills to support a mission or a social cause” (McDonough Dolmaya, 2011). However, in case of crowdsourcing the biggest benefit is the sense of community, therefore such companies as *Facebook* appeal to this when seeking for volunteers who would help translate their website.

Another important thing that is thought to be an issue is translation visibility. Though crowdsourcing initiatives are helping translation become a more visible practice through publicizing the activities of volunteer translators, visibility alone will not make translation seen as a high-status profession. This can only be done when people actually faces the difficulties in translation and raise them publicly. Then fan translators may require more practical experience and only in this way the visibility of this profession can actually gain more attention. Other initiatives, like *Facebook*, allow the community to make it responsible for quality control regardless of whether the users have any formal training in translation. However, similar activities do not do not emphasize the difficulty of the translation process, therefore public perception of translation may be lowered, since it may seem as if any bilingual can easily complete the task. All these factors may assume that translation is more suitable for a hobby than a profession.

The last issue is with less popular languages, or, in other words, with those making not such a big impact on the market. Usually markets with a small number of potential customers will get less



translated content than areas where the languages are spoken in several markets. Even though it is pointed out that crowdsourcing may help languages with minor number of speakers take somewhere near the leading positions, languages with more speakers online will have a larger pool of volunteers to choose from, therefore translations in Spanish, French are likely to be finished sooner than the ones in Romanian or Swahili, despite the fact that the source text is made available at the same time (McDonough Dolmaya, 2011).

Having discussed the most important aspects of website localization and crowdsourcing further on the selected social networking sites translated by means of crowdsourcing and by a translation agency are going to be analysed and compared on textual, linguistic and pragmatic levels.

## 2. PRACTICAL ANALYSIS OF *FACEBOOK* AND *GOOGLE+*

Having discussed peculiarities of localization and crowdsourcing it is a high time to check these theories in practice. In the following sub-chapters the two selected websites are going to be analysed and their localization quality checked. After that results of investigation are going to be discussed and conclusions given.

### 2.1. Methodology

In this chapter English and Lithuanian versions of menus and elements comprising user interfaces of social networking sites, such as *Facebook* and *Google+*, are going to be analysed. The most prominent errors are going to be provided and described and quality evaluation given as well. During the analysis of the results comparative, content and descriptive analysis methods are going to be implemented. For evaluation of quality of translation the scheme proposed by Miguel A. Jimenez-Crespo *The different components of web localization quality* is going to be used. The object of the research is going to be various textual elements containing user interface and menus. Then those elements are going to be grouped and analysed according to textual, linguistic and pragmatic features and the evaluation of translation quality is going to be given.

### 2.2. Comparative analysis of *Facebook* and *Google+*

*Facebook* is one of the largest social media site having roughly about 2 billion active users. Moreover, this site is being translated by its users via a special *Translate Facebook* page, accessible to anyone who wants to contribute to translations in their native language. Once accessed users see different strings provided and can start translating at once or review other users' translations and edit them. Which translations appear on the site and which do not depends on the number of votes they get from other users. Translation of a string with the highest number of votes soon appears on the site, but other versions of the same string are also visible, therefore the proposed translations can be changed at any time, if needed. All contributors can see their number of contributions as well as how many votes they got. *Google+*, on the other hand, has approximately about 375,000,000 active members and unlike *Facebook* this site has been translated by a professional translation agency. All the focus in the following subchapters is going to be devoted to the translation of both sites, however *Google+* in this work is treated as an example of a well localized website.

#### 2.2.1 Textual errors

Since *Facebook* is translated by volunteers, it seems they do not devote so much attention to unified translations of various items. Consistency can only be guaranteed if some specific term has

been added to the glossary. Otherwise users themselves have to remember their proposed versions and use them as required. However, volunteers' community is quite large, therefore the task becomes much more difficult as they tend to use whatever comes into their heads and even without pre-checking of the meaning of similar words. As a result translations become inconsistent and may impair the communication process.

Below there is a table containing various types of inconsistency found on *Facebook*.

**Table 1.** The most prominent errors of inconsistency on *Facebook* social website

No.	English	Lithuanian	Issue
1.	Events / Event	Įvykiai / Renginys	Inconsistent translation of 'event'
2.	Top Comments / Top Comments (unfiltered)	Populiariausi komentarai / Geriausi komentarai (nefiltruoti)	Inconsistent translation of 'top comments'
3.	New Message	Nauja žinutė / Naujas pranešimas	Inconsistent translation of 'message'
4.	SHORTCUTS	Nuorodos / Greitosios nuorodos	Inconsistent translation of 'shortcuts'
5.	Who can post on your timeline? / ...they appear on your Timeline? / your own posts	Kas gali skelbti jūsų metraštyje? / ...prieš parodant juos tavo laiko juostoje? / prie Jūsų įrašų	Inconsistent translation of 'your'
6.	App	Programa / programėlė	Inconsistent translation of 'app'
7.	Share	Dalintis / Bendrinti	Inconsistent translation of 'share'
8.	Like	Patinka / mėgsta	Inconsistent translation of 'like'
9.	Edit Profile / Edit Preferences	Redaguoti profilį / Keisti nustatymus	Inconsistent translation of 'edit'
10.	Block users	Blokuoti asmenys / Blokuoti naudotojus	Inconsistent translation of 'block' and 'users'

Source: created by the author

One of those inconsistently translated terms is the word *event* (1). In Lithuanian this noun is polysemous and several meanings prevail. This word appears on the home page as a menu item in two places. In the first case it appears in plural form and has been translated as *Įvykiai*, but just below that menu the same item is displayed as *Renginys*. Translators of *Google+* in this situation have chosen to stick to the version of *įvykis* and use it everywhere, thus it may seem such translation should fit to *Facebook* as well, but actually it is not. In the case of *Facebook* the meaning of *event* can be found in the glossary where it is provided as *renginys* and so this version must be used in any case of its appearance. Since under the *Event* option both versions are met, it becomes clear that a user who translated this word did not pay attention to the glossary meaning, as well as other users, otherwise such situation would not be present. Translation of *Event* is not the only case when glossary terms have

been ignored. Other the most prominent occurrences include translation of *app* (6) and *edit* (9). The former word in some places is translated as ‘programa’, while in some other areas term ‘programėlė’ appears and this inconsistency continues throughout the website, even though the glossary meaning is ‘programėlė’. There can be many reasons as to why the glossary has been ignored, but it can be assumed that such a popular word these days like ‘an app’ is so well-known there is simply no need checking it – everyone knows what it means. Since it is very difficult to say who was responsible for rendering this word, it might have been this term was translated by a beginner who is not familiar with the process of translation. However, there is also a high chance of possibility other users might not even know such thing as a glossary even exist. Moreover, that glossary page does not contain a search field and so users have to use their browsers’ search function which, to some extent, is not very comfortable and requires some time in finding a certain word. The second word *edit* in the group options menu used to be translated as *Keisti nustatymus* and in the profile settings as *Redaguoti profilį*. However, few days ago at the time this work being written translation in the group options menu has been changed to *Redaguoti nustatymus* as it should be. This fact only proves that a website translated by means of crowdsourcing is like a living organism and translations can be changed at any point. This also means inconsistencies do care for voluntarily working *Facebook* translators as well as quality in general.

The same situation was noticed in translation of *Top Comments* (2). Earlier this feature for sorting comments was translated like *Populiariausi komentarai* (the most popular comments) and *Geriausi komentarai* (the best comments). On the day of writing this work it is already translated as *Aktualiausi komentarai* (the most relevant comments) in both cases of their appearance. In fact such translation a bit distorts the meaning, since by ‘top’ mostly liked comments are meant and the ones which are somehow the most meaningful below that post.

Another inconsistency involves a word ‘message’ (3). This word is not found in the glossary, thus it can be translated both as ‘pranešimas’ or ‘žinutė’. In the messages menu dominates translation of ‘žinutė’, but whenever one clicks on the new message button a separate tab opens at the bottom for typing the name of an addressee and text itself. Here for some reason the tab is called *Naujas pranešimas*. In the target language meaning between these two words is not very different, thus these words are often used interchangeably. Yet still one word had to be selected in translation and in this context ‘žinutė’ seems to be more suitable meaning as usually it is a very short text carrying some kind of information, just like the meaning in Lithuanian dictionary states<sup>2</sup>.

Very similar case is also with a share button – probably one of the most clickable buttons on *Facebook*. Translation of this button is not indicated in the glossary, thus various versions seemed to

---

<sup>2</sup>[http://lkiis.lki.lt/paieska?header=%C5%BEinut%C4%97&resourceId=207&remoteRecordId=12669594&p\\_p\\_id=LKISearch\\_WAR\\_LKISearchportlet](http://lkiis.lki.lt/paieska?header=%C5%BEinut%C4%97&resourceId=207&remoteRecordId=12669594&p_p_id=LKISearch_WAR_LKISearchportlet)

appear more easily. In the home page this button is translated as *Dalintis* and appears at the bottom of every post visible there causing any problems to users. But having entered into any group, the same button displayed on the upper part of the group window and asking users to share the group is shown with a command *Bendrinti* and that is clearly an error of inconsistency. Moreover, very interesting thing happens when hovering over this button with a cursor. Then a description of the command appears and informs the user saying *Pasidalinkite šia grupę*. Hence, in the very same place two different translations of share command appears. Nevertheless, these words in Lithuanian can be used as synonyms, thus website usability should not be impaired very much. *Google+* in this regard also uses share button and in Lithuanian is sounds as *Bendrinti* or *bendrinama* depending on situation. In addition to this, *Google* decided to use another method and replaced like, share and comment buttons for respective icons denoting these functions.

Another inconsistency is found in case of like button. Actually all three mostly used buttons, i.e. like, comment and share, are provided inline and so their translations should also be provided in the same manner – by means of infinitive. However, only *Comment* and *Share* are conveyed in the target language as infinitive verbs, while *Like* is translated as *Patinka*, which is a verb of the present tense. Furthermore, when the same word appears in different situations it is not translated as *patinka* anymore, but is substituted for another verb *mėgti*. For example, when a user likes a post, comment or something else the segment goes like ... *liked this* which in the target language is rendered as ... *tai mėgsta*. Not only is the word *Like* translated differently, but the tense used is different as well. The described issue in this case can be solved by choosing either *patikti* or *pamėgti* and using these throughout the website. In fact, translating *Like* as *Pamėgti* would sound even more natural.

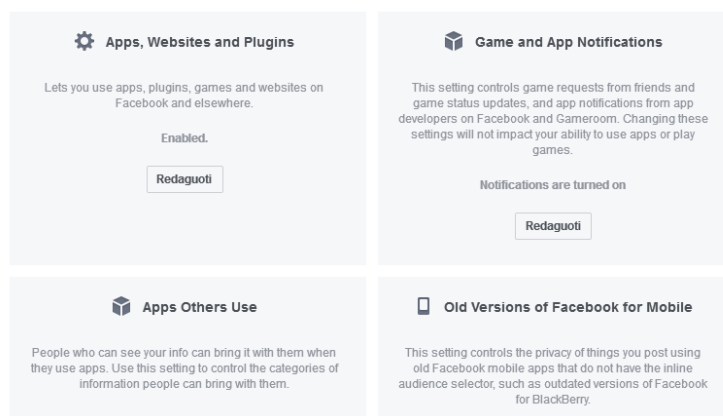
Inconsistency regarding tone also can be found. As it was mentioned in the theoretical part, social networking sites are expected to represent more informal tone rather than formal. Formality in Lithuanian is created by using third person plural pronouns along with respective verbs. However, when the website is being translated by so many different users sometimes tone gets mixed like in this example: *Who can post on your timeline? – Kas gali skelbti jūsu metraštyje?* Then, right below this string comes another one asking the user whether he or she wants to review any posts that friends tag them in before their appearance on the timeline. In this case the third person plural pronoun becomes the second pronoun singular in the target language and thus signifies different level of formality: *Peržiūrėti įrašus, kuriuose draugai tave pažymi, prieš parodant juos tavo laiko juostoje?* Since both of these segments appear in the same place just below one another, communication between the user and the website suddenly turns into very odd process. In addition to this, some other segments in the same Timeline and Settings page even include very formal writing of pronouns, such as *Jūsu* which are more expected to be met in some business letters or legal content. English pronoun *your* does not clearly signify to the target language speaker what level of formality it belongs to as it may be used in

both ways. Furthermore, no style guide is prepared for Lithuanian which already impairs the task and referring to style guides of other countries would be of no use as different guides require different tone. However, putting a pronoun in the capital letter is not a good strategy as well. In this case at least the third person plural pronoun should have been used which is a normal practice in Lithuanian translations. Moreover, *Google+* localizers also rendered pronouns *your* as *jūsų* and only in some places other pronouns like *mano* (mine) signifying informal tone were used. Nonetheless in such case tone in the target language becomes more formal rather than informal, but as different style guides have shown formality mostly depends on the culture of that locale.

Translation of ‘shortcuts’ (4) seemed to cause problems as well. In IT context this term is usually translated as ‘nuoroda’ without highlighting the feature of possibility to reach some content in a more quicker way. In *Facebook* this term stands as a title of the menu where a user can reach his mostly visited groups by one click, thus translation *Greitosios nuorodos* fits very well. Nevertheless, other crowdsourcers did not devote so much attention to it and in other areas rendered it simply as *nuorodos*. Difference in this case is not so big, thus it may be assumed usability should not be affected much. However, translation still needs to be unified throughout the website and the existing version changed to all capitals as it is in the source.

Another peculiar inconsistency is translation of titles of an action. In English the same word, for example ‘block’ (10), can be used as a noun and as a verb, thus translator must find out which meaning the word takes in various contexts. During translation it is usually indicated in the description whether it is a title or something else. If it is a title, it should be translated as a noun. In this case ‘block’ in both places of its appearance denotes an action and so they should have been translated as ‘blokavimas’, i.e. as a noun denoting this particular action. Also, putting ‘users’ as ‘asmenys’ (persons) is not a good strategy, since people in the social networking sites do not communicate as live persons, but rather using their profiles. Thus, such translation is also a crude mistranslation, while *Block users* should have been translated as *Naudotojų blokavimas*.

Yet the most serious issue on the textual level is untranslated text. There are only one or two untranslated units on the homepage where users get first after logging in, but going deeply into any menus chance of meeting a foreign language text becomes higher. This especially concerns Settings, Help or Advert sections.



**Figure 2.** Entire blocks of untranslated text

All these sections include specific terminology and consist of much more text thus requiring more time and effort to complete. Since there are no deadlines provided and users choose the content they want to translate on their free will, these pieces including foreign text units can be left untouched for an indefinite period of time. Such situation can cause serious usability problems as users would not be able to get help in their own language. Also, Lithuanian version of *Facebook* is full of half-translated content, for instance: *Kampanija: Choose your objective, Įjungti pokalbius to see who's available, Comment rating is Išjungta* and so on. Furthermore, there are half translated menus, untranslated pieces of text next to translated strings and similar uncompleted content. Appearance of half-translated sentences can be easily explained. Usually strings on *Facebook* translation page appear as full sentences, but sometimes one or another string can be split into two parts and if translator is not attentive enough, the following string is simply skipped assuming it is another separate sentence, especially when that second part contains a bit complex structure or unknown words. In case of *Google+*, such textual errors are not present and the whole website seems to have been fully localized indicating the certain level of quality.

In this subchapter only the most peculiar cases of textual errors have been discussed. Having taken into consideration issues described above it is possible to state that the localized version of *Facebook* is full of inconsistencies, untranslated or partially translated text throughout the site, thus textual quality is satisfactory only and in some cases this may cause usability problems for the users. Quality can be improved by adding to the translation platform additional functions which would inform about inconsistent translations or untranslated pieces.

### 2.2.2 Linguistic errors

Another group that corresponds to quality assessment is linguistics. It has already been proved that errors do appear in the crowdsourced translation on the textual level and linguistics is no exception. One of the most prominent issues on this level are case usage errors, especially when it

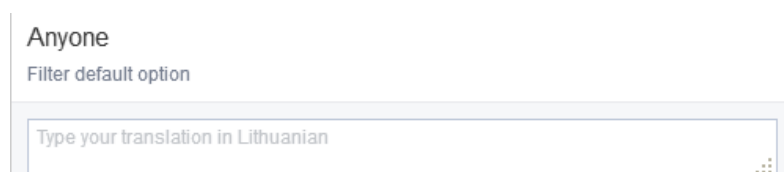
comes to numbers. During the process of translation numbers are being replaced by special tags, called variables, which will later represent different numbers like in this instance: *{number} mutual friend*. The part in curly brackets is not translatable and only signifies what that variable stands for and in future it will be replaced by a number. Thus in the target text the whole string should look like, for example, *1 bendras draugas*. However, the given example in the target text is as simple as it can be and yet far from perfect. First if all, example provided above denotes only the masculine gender and in order to avoid this problem letter *-a* should be added in parenthesis after *bendras* and *-ė* after *draugas*. Secondly, such method would be useful only with numbers 1, 21, 101, 1001 and so on. When other numbers appear inflections should be changed accordingly, but the system itself cannot do that, therefore translator would need to add any possible inflections in parenthesis making the string considerably long and quite complex to read. One of the best ways in this case would be to put that variable at the back of string, for example: *Bendru draugu: 1*. Even though the original sentence is in singular, comment below that string indicates that this line will be used to inform about a number of common friends, therefore it should include plural as well. Such version would solve issues related with gender and inflections and keep almost the same number of characters.

The example provided earlier was given only to illustrate, how volunteers deal with numbers as they appear in the text. Usually *Facebook* provides additional segments for number-related content and translators are allowed to enter different translations for strings containing numbers like 1, 2, 21, 31 or 10, 50, 100, etc. Nonetheless, errors still tend to appear and translation of *9 Comments* becomes *9 Komentaru*. Such string is displayed below every post, next to the total number of likes or different emoticons denoting number of emotions expressed. Presumably the source text should look like *{number} Comments*, but the one who translated it did not pay particular attention to the inflection of the word *komentarai* and to writing of capital letters. Unfortunately *Facebook* translation page does not include search function and it is not possible to check, how actually the source text might have looked like. If there were not any other strings for distinct types of numbers provided, then it should have been translated as *Komentaru: 9*. However, if additional segments were provided, then the translator might not referred to contextual information about that string or put it in a wrong way. The very same issue is found with strings denoting, how many users along with your friends liked or expressed emotions on various posts. In English it goes like *(username) and 21 others*, while in Lithuanian it is *(username) ir 21 kiti (-u)*. Such translation using parenthesis sounds quite odd to the target language speaker and does not seem to fit right, so it has to be edited to something like *(username) ir dar 21*. That is only one of possible suggestions when the pronoun *others* is being replaced for the adverb *dar* (more). *Google+* in this regard does not use similar phrases and the number of total comments is provided next to the icon of likes. This strategy is very useful, as it allows avoiding similar issues. Other analogous cases include such translations as *pasidalino (username)*



*įrašas* or *mėgsta tavo įrašas jos metraštyje*, which seems to have appeared simply due to lack of attentiveness.

One more similar error is found in Settings section, where the user is asked to choose who can send him or her friend requests and the setting chosen is *Everyone*. In Lithuanian these settings are translated as *Kas gali Jums siųsti kvietimus draugauti? – Visiems*. Translation of ‘everyone’ as ‘visiems’ (for all) sounds awkward and incomprehensible as the dative case is used instead of nominative (*visi*). Such situation may have arisen due to two reasons. First – the system took some earlier translation of ‘everyone’ from different context and put in this place or, the second one, translator tried guessing the situation in which this string should appear and applied the possible case referring to the explanation provided about that string. Even though it may seem as a simple mistake, translation of units without and additional words around them impairs the task very much. Sometimes even an explanation given does not contribute to trying to find a connection between the previous or following strings, since very often words are taken from various places and are provided at random. The picture below illustrates similar situations very well.



**Figure 3.** String ready for translation on *Facebook* localization page

As it is seen, only one word is provided and all that the translator knows it is some kind of filter option, thus it is still not clear in which section that filter is going to appear and what kind of pronoun would fit best, since in the target language it can be rendered as *nė vienas / nė viena, bet kuris / bet kuri, bet kas* or *niekas*. In such cases translators have to use their best judgement hoping they have made the right decision.

Misleading or slightly changing the meaning translations make quite a big part, thus the quality is diminished even more. Minor meaning shifts includes such cases as *Friend lists* is translated as *Draugų sąrašas*, even though a plural form *sąrašai* should be used there, or *More contacts* is translated as *Daugiau draugų* instead of *Daugiau kontaktų*. Nevertheless, there are more serious instances. In the group menu there is an option *Unpin from top*. Having selected this option the group is no more shown at the top of other groups. However, translation of this command sounds like *Nebeteikti pirmenybės žymei* (do not prefer this tag) and seriously distorts original meaning, thus impairing user usability as well. It looks like the line appeared in this place by a mistake, since it does not have anything in common with the source text. This segment should have been translated something like *Atžymėti iš viršutinės pozicijos* or *Nebeteikti viršuje* (do not show at the top). The latter is a bit more distant translation from the source, but it clearly denotes, what that function is for without affecting the site

usability. Yet this is not the only example from the very same menu. Just below this line comes another one, which in the source is *Hide from shortcuts*. Even though it looks like a simple string, it still was not that easy for the translator who translated it as *Paslėpti greitųjų nuorodų sąrankoje*. For some reason the one who was responsible for this line translated ‘shortcuts menu’ as a ‘setup’, which is incomparably different type of thing. To some extent, menu can be called as being a set or a collection of items, as one of the meanings that Lithuanian words possess, but in this context *sąranka* is just not an appropriate word, because it signifies the process of installing an application to a computer and here it clearly denotes shortcuts menu, not a setup. Moreover, translation involves spelling error and *greitųjų* suddenly becomes *greitųjų* that distorts the meaning even more. All in all, this string should be translated as *Paslėpti greitųjų nuorodų meniu*.

Another case is translation of ‘tabs’. Due to some reason it has been translated as ‘langas’ (window). To some extent it has some features of a normal program window, but in *Facebook* it is a way much smaller item meant for messaging and it cannot be treated as a window, therefore ‘skirtukas’ would fit best in this context. Translation of ‘following’ is also problematic. When a user hits ‘Follow’ button he or she can see another user’s comments, likes and shares without becoming friends with that user. This function in Lithuanian is conveyed pretty literally as *sekti* and after clicking it the user becomes *sekamas*. Such translation sounds at least very odd to the target language speaker and may seem that others from the moment of clicking it would see one’s every little move on the Internet. Nonetheless, the biggest problem is with companies or organizations. Their profile pages also have this button and after clicking on their status changes to *Prenumeruojama* (subscribed), not *sekamas*. Moreover, in the area where it gives the number of how many people follows that page this word is already translated as *Stebi*. Thus, on the localized *Facebook* website several different translations of ‘follow’ (including different parts of speech) appear: ‘sekti’, ‘sekamas’, ‘stebi’, ‘prenumeruojama’. The latter is absolute mistranslation and has nothing in common with the source meaning. What regards *Google+*, this website has such function as well, the only difference is that the localizers stuck to one meaning – ‘stebėti’ and incorporated this word throughout the site: ‘stebima’, ‘pradedama stebėti’, ‘stebėtojai’ and so on. Hence, this only proves that word-for-word rendering of ‘follow’ is not the best solution possible, while translating it as ‘prenumeruojama’ would not be acceptable at all. Taken all these peculiarities into consideration it is possible to claim that the existing translation on *Facebook* should be changed to ‘stebėti’.

It has been noticed that when it comes to an unknown word, which denotes some function or an item, volunteers tend to translate it according to the result shown after having it pressed or a graphical representation of an element like it was in case of a ‘tab’. However, sometimes this strategy is not as bad as it may look like. Not only *Facebook* provides a social networking site or a platform for voluntarily proposed translations, but it also offers a possibility to become a part of the company’s

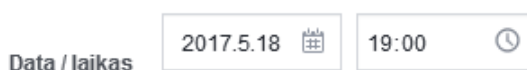
team. For this reason there is a separate section with job offerings. Users can get to this section by clicking *Careers* button. Surprisingly, in the target language it is not translated as *Karjera*, but as *Darbo pasiūlymai*, i.e. ‘job offerings’. Even though such translation is far from semantic meaning of the source word, it very well presents what one can expect after clicking on it and does not seem to affect user usability at all. However, such a freedom in translation should not be tolerated in cases when there is not any actual reason for that.

Some issues also involve localization of calendar and date. Calendar in *Facebook* is used when creating an event for picking the exact start and end dates.

Gegužės 2017						
Pir	An	Tre	Ket	Pen	Šeš	Sek
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

**Figure 4.** Localized calendar on *Facebook* website

At the top of the calendar current month and year are provided, but the localized version is missing abbreviations, such as ‘m.’ and ‘mėn.’ (‘m.’ stands for year and ‘mėn.’ stands for month) – *Gegužės 2017*. Moreover, in Lithuanian year should be taken an initial place proceeded by the month, name of the month should start in lowercase and a case from the genitive must be changed to the nominative. Another problem is with abbreviations of days which in this calendar are abbreviated to three or two characters. According to the State Commission of the Lithuanian Language, names of days are to be abbreviated to only one letter, for example: pirmadienis (Monday) – P, antradienis (Tuesday) – A, penktadienis (Friday) – Pn and so on<sup>3</sup>. Having chosen a day the calendar disappears and set dates are shown in a normal date format. However, in the target language version date constituents are separated by dots, not by short dashes as it should be: YYYY.MM.DD.

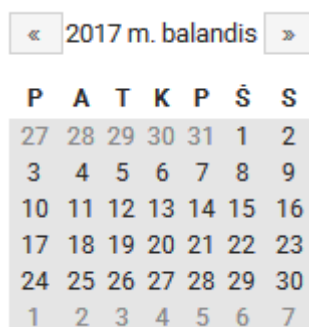


**Figure 5.** Localized date and time format on *Facebook* website

At first sight it seems this problem may be related with technical peculiarities of *Facebook*, than being a translator’s mistake, but in English the order of date constituents is adapted to that locale and slashes are used as separators. This fact proves the date is localizable, but the one who did that had chosen wrong separators and the date format must be changed now to YYYY-MM-DD format.

<sup>3</sup> <http://www.vlkk.lt/konsultacijos/10807-savaites-dienos-santrumpos-sutrupinimai-simboliai>

Google+ localizers in this respect did everything right. Date order has been adapted to the target language needs, correct separators have been used and names of days have been abbreviated to only one symbol as required.



**Figure 6.** Localized calendar on Google+ website

Other errors include writing of capital letters, usage of slash or missing gaps, for instance: *Nuotraukos/Vaizdo įrašai*, *Būsena/Veikla*, *Pranešti apie Grupę*, *11 881 sekėjų(-ai)*. According to Lithuanian language norms two words separated by the slash must include a gap before and after that slash and initial letter of the second word should be lowercased as both expressions belong to the same sentence. Exceptions are applicable only for units, such as km/h, g/cm<sup>3</sup>, m/s, etc. Moreover, gaps are to be used before parenthesis as well, while capital letters are used only for names, titles or brands, thus translation given above should be changed as follows: *Nuotraukos / vaizdo įrašai*, *Būsena / veikla*, *Pranešti apie grupę*, *11 881 sekėjų (-ai)*. Also, in some minor cases wrong quotation marks have been used, for instance *Siųsti per “Messenger”*. In the target language the opening quotation mark is positioned at the bottom and the closing quotation mark at the top, therefore the aforementioned string should be changed to *Siųsti per „Messenger“*. In some other few instances quotation marks have not been used at all and even the name *Facebook* was inflected by adding an -e ending, which denotes the locative case in Lithuanian, for example: *Dabar galite išsaugoti Facebooke esančius dalykus vėlesniam laikui*. Very often such frequently used words become an integral part of culture in any linguistic community and are adapted according their pronunciation. However, these adapted words are only allowed to be used in a colloquial context and *Facebook* is definitely not the best place for that. It seems that in the given example another was strategy used, when an inflection is added to foreign surnames in order to facilitate communication process with a target language speaker. Nonetheless, inflections are not allowed to be added to such brand names, thus they have to be left as is and quotations marks around them are to be used.

Another constituent of an internal quality, linguistic peculiarities, proved to include some serious errors that cannot be ignored. Some serious mistranslations and word-for-word translations have been found as well as translations that do not comply with the target language linguistic norms. All in all,

linguistic quality also proved to be satisfactory only and still quite a lot of work needs to be done on this level.

### 2.2.3 Pragmatic errors

The biggest issue in the localized version of *Facebook* in terms of pragmatics is inconsistent usage of pronouns as it was seen earlier. Such situation creates different level of politeness which is sometimes even more exaggerated by using a pronominal form written from the capital letter, e.g.: *Kai ką nors užblokuojate, tas asmuo nebegali matyti dalykų, kuriuos Jūs talpinate savo laiko juostoje, žymėti Jus nuotraukose, kviesti Jus į renginius ar grupes, užmegzti su Jumis pokalbį ar pakviesti Jus draugauti*. In Lithuanian pronouns ‘jūs’ (your) and ‘jūsų’ (yours) are very often used in translation of software and websites and this already changes the tone of a product into quite a formal one. Additional usage of these polite forms of pronouns turns social networking site into very serious type of a website and increases distance, though such site is expected to be more informal. Such level of formality would be more likely to be found in business type of websites or social networks like *LinkedIn*. Moreover, different level of politeness in one place poses a threat to user’s faces. When it seems that their expectations are going to be met, suddenly informal pronoun ‘tu’ appears and social distance becomes much closer while the user is left confused.

*Google+* has also been translated in quite a formal way. That dominance of the third person plural pronouns turns the site into some other type of website different than social networking site, e.g.: *Rodyti žmones, kurie pridėjo jus prie draugų ratų, Kas gali siųsti jums pranešimus, Kas gali komentuoti jūsų viešus įrašus?* However, such is nature of the Lithuanian language. Using only the second person singular pronouns the distance would be minimized to minimum, but the tone of the site would be too informal. Since it is not known who and when will be using the site the informal pronouns only would not fit as users of higher social statuses may feel offended.

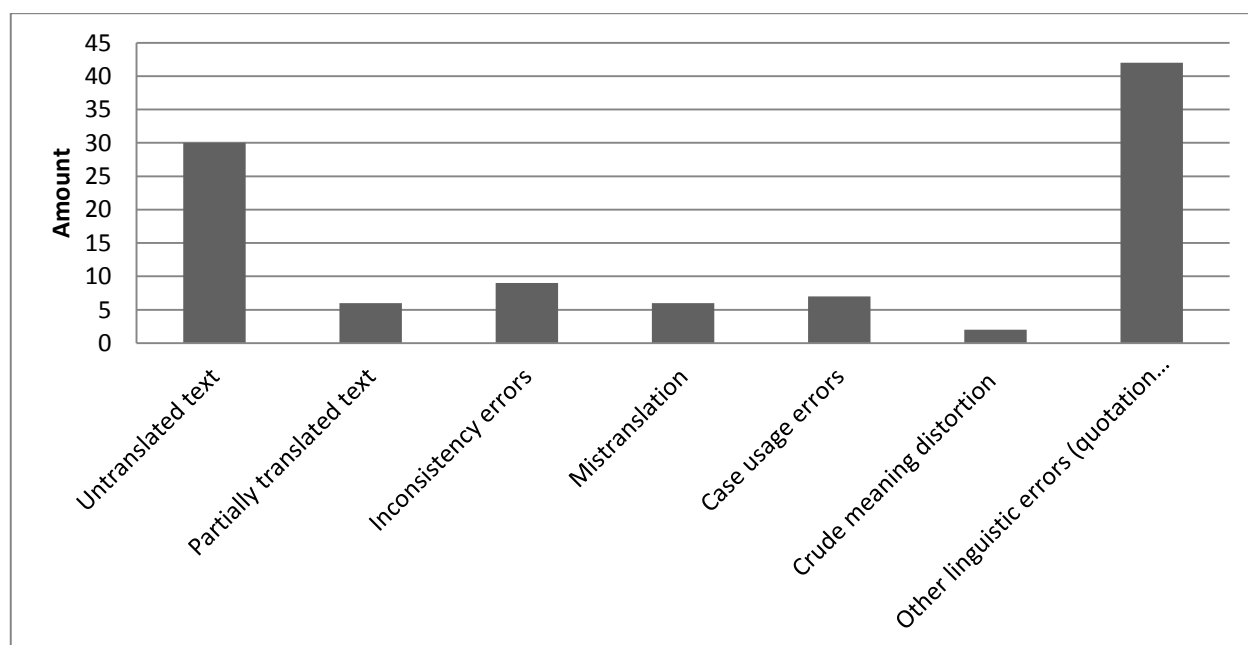
### 2.3. External quality: web usability, functionality and compliance with client’s commission

As it is presented in the quality assessment scheme, web usability and functionality to some part depends on the intrinsic components of internal quality. Functionality is less dependent from the latter as this factor deals with technical properties of a website rather than linguistic ones. Various hyperlinks or menus may still work regardless of the quality of translation, if functionality is flawless. Web usability is a more decisive component, since textual or linguistic errors may leave bad impression about a site or Internet in general. However, some serious mistranslations do not seem to affect usage of *Facebook* between Lithuanian language users much and it still remains the most

popular social networking site<sup>4</sup>. Since *Facebook* allows choosing other languages, it is difficult to say how many users browse in their native language and how many of them tend to tune in the other one.

In terms of compliance with client's commission it is also very difficult to say how this factor should be evaluated. Since this is relatively new kind of subject, it lacks the input of scholars, thus there are no established criteria for such evaluation and assessment can be done only on a very subjective level. Nonetheless, having done the analysis it may be assumed that users, as clients to some extent (those who use *Facebook* not only for their social needs, but contribute to localization of it, vote for other users' translations or notifies about wrong translation in the special group) should not be happy about it and would give quite a low score. On the other hand there are no style guides that would help a lot and this also obscures the whole process. Moreover, while some volunteers try to do their best in translations, others write whatever comes to their heads just for the sake of fun or trying to make out. Thus, translation of *Facebook* can be seen as a good starting point for beginners who want to gain some experience and knowledge about the whole process of localization. What regards *Google+* is the difference in quality which is seen very well between these two sites. It may be assumed that *Google* should be very happy about the Lithuanian version of their social site as errors are quite difficult to be found, which testimonies how careful work was done in every step of localization process.

#### 2.4. Results of investigation



**Figure 7.** The distribution of errors in localization of *Facebook*

<sup>4</sup> <http://www.tns.lt/lt/news/tns-lt-moterys-dazniau-naudojasi-facebook-%2C-vyrai-youtube/>

In total 102 examples have been analysed. The most prominent errors proved to be various language-related errors, such as addition of unnecessary gaps, missing gaps, usage of capital letters in the middle of text and so on. Untranslated text made up around 29% of all the cases which is quite big part as well. In total only 7 instances of case usage errors has been found, but this is the third biggest amount. Partially translated text did not make such a huge part as expected and only 6 cases were found. Mistranslations also covered 6 cases and crude meaning distortion has been found in 2 cases only. Despite that a localized website should not contain any meaning distortion errors as even the very small amount of them may result in serious consequences.

All in all such analysis seems to be too small and it still needs to be done in more broadly and deeply way. *Facebook* contains lots of strings, but there is not any method how to extract and analyse all or at least only the bigger part of them.

## CONCLUSIONS

From the theoretical and practical parts the following conclusions can be drawn up:

1. The set out criteria proved to be useful only for evaluation of an internal quality.
2. The website localized by means of crowdsourcing proved to contain some serious errors in translation, quite big parts of untranslated text and cases of wrong translation.
3. Having compared websites translated in two different ways it has been found that quality of the website translated by means of crowdsourcing was way much lower that the one's that has been translated by the professional translation agency.



## REFERENCES

- Archer, D., Aijmer, K., & Wichmann, A. (2012). *Pragmatics: An Advanced Resource Book for Students*. New York and London: Routledge
- Aris, H. & Din, M. M. (2016). Crowdsourcing evolution: Towards a taxonomy of crowdsourcing initiatives. *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, 1-6.
- Baer, N. (2010). Crowdsourcing: Outrage or opportunity? *Translorial: Journal of the Northern California Translators Association*. Retrieved from <http://translorial.com/2010/02/01/crowdsourcing-outrage-or-opportunity>
- Brabham, D. C. (2008). Crowdsourcing as a model for problem solving: An introduction and cases. *Convergence: The International Journal of Research into New Media Technologies* 14(1), 75-90. doi:10.1177/1354856507084420
- Brabham, D. C. (2013). *Crowdsourcing*. Cambridge, Massachusetts; London, England: The MIT Press.
- Chroust, G. (2007). *Software like a Courteous Butler-Issues of Localisation under Cultural Diversity*. Austria. Kepler University
- Cronin, M. (2010). The translation crowd. *Revista tradumàtica* 8. Retrieved from <http://www.fti.uab.cat/tradumatica/revista/num8/articles/04/04.pdf>
- European Commission. (2012). *Crowdsourcing Translation*. Luxembourg: Publications Office of the European Union
- Flanagan, M. (2016). Cause for concern? Attitudes towards translation crowdsourcing in professional translators' blogs. *The Journal of Specialised Translation* 25, 149-173, [http://www.jostrans.org/issue25/art\\_flanagan.pdf](http://www.jostrans.org/issue25/art_flanagan.pdf).
- Gibb, F. & Matthaiakis, I. (2007) A framework for assessing web site localisation. *The Electronic Library*, Vol. 25 Issue: 6, pp. 664-678, doi: 10.1108/02640470710837100
- Huberman, B. (2008). Crowdsourcing and attention. *Computer*, 41(11): 103–105.
- Janoschka, A. (2003). *Web Advertising*. Amsterdam-Philadelphia: John Benjamins
- Jeney, C. (2007). *Writing for the Web: A Practical Guide*. Columbus: Pearson Prentice Hall
- Jiménez-Crespo, Miguel A. (2013). *Translation and Web Localisation*. London, New York: Routledge
- LoCastro, V. (2012). *Pragmatics for Language Educators: A Sociolinguistic Perspective*. New York and London: Routledge
- McDonough Dolmaya, J. (2011). The ethics of crowdsourcing. *Community Translation 2.0* Nr. 10.
- O'Hagan, M. (2009). Evolution of User-generated Translation: Fansubs, Translation Hacking and Crowdsourcing. *Journal of Internationalisation and Localisation* 1, 94-121.

- O'Hagan, M. (2016). Massively Open Translation: Unpacking the Relationship Between Technology and Translation in the 21st Century. *International Journal of Communication* 10, 929–946.
- Perrino, S. (2009). User-generated Translation: The future of translation in a Web 2.0 environment. *The Journal of Specialised Translation* 12, 55–78, [www.jostrans.org/issue12/art\\_perrino.pdf](http://www.jostrans.org/issue12/art_perrino.pdf).
- Pym, A. (2010a). Translation research terms: a tentative glossary for moments of perplexity and dispute. *Translation Research Projects* 3. Ed. Anthony Pym. Tarragona: Intercultural Studies Group, 75-110.
- Pym, A. (2010b). *Website Localization*. London, New York. Routledge
- Rajala, R., Westerlund M., Vuori, M., Jukka-Pekka, H. (2013). From Idea Crowdsourcing to Managing User Knowledge. *Technology Innovation Management Review* 3(12): 23-31. <http://timreview.ca/article/750>
- Sandrini, P. (2008). Localisation and Translation. *Acts of the LSP Translation Scenarios (MuTra)*. Conference proceedings. Vienna.
- Surowiecki, J. (2004). *The Wisdom of Crowds*. London: Abacus.
- Zaidan, O. F. & Callison-Burch, C. (2011). Crowdsourcing translation: Professional quality from nonprofessionals. *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies* (Vol. 1). Association for Computational Linguistics.