

The Journal of Internationalisation and Localisation
Volume I
2009



Lessius University College/K.U.Leuven
Department of Applied Language Studies
Sint-Andriesstraat 2, 2000 Antwerpen

Localization Industry Standards Association (LISA)
Domaine en Praël CH-1323 Romainmôtier, Switzerland

Brigham Young University
Department of Linguistics and English Language
Provo, UT 84602, USA

The Journal of Internationalisation and Localisation
Volume I
2009

© 2009 by

Lessius University College/K.U.Leuven
Department of Applied Language Studies
Sint-Andriesstraat 2, 2000 Antwerpen

Localization Industry Standards Association (LISA)
Domaine en Praël CH-1323 Romainmôtier, Switzerland

Brigham Young University
Department of Linguistics and English Language
Provo, UT 84602, USA

Editors

Hendrik J. Kockaert (Lessius University College/K.U.Leuven)
Arle Lommel (Localization Industry Standards Association)

Editorial Board

Dr. Esperanza Alarcón Navío, Facultad de Traducción y Interpretación, Universidad de Granada,

Spain

Michael Anobile, *Managing Director*, Localisation Industry Standards Association [LISA],

Romainmôtier, Switzerland

Prof. Dr. Bassey E. Antia, Department of Languages and Linguistics, University of Maiduguri,

Borno State, Nigeria

Patricia Egan, Technical Communication, University of California at Berkeley Extension, USA

Jorge Estevez, *Group Operation & Processes Manager*, Jonckers Translation & Engineering,

Brussels, Belgium

Dr. Miguel A. Jiménez-Crespo, Department of Spanish and Portuguese, Rutgers, The State

University of New Jersey, USA

Barbara Inge Karsch, *Terminology Researcher*, Microsoft Corporation, Redmond, USA

Rolf Klischewski, *Games Localisation Consultant*, Germany

Dr. Hendrik J. Kockaert, Lessius University College/K.U.Leuven University, Belgium

István Lengyel, *COO*, Kilgray Translation Technologies, Hungary

Dr. Arle Lommel, *Publications Manager*, LISA, USA

Julia Makoushina, *Co-Owner and Operations Director*, Palex Languages and Software, Tomsk,

Russia

Prof. Dr. Alan K. Melby, Linguistics and English Language, Brigham Young University, Provo,

USA

Dr. Sharon O'Brien, School of Applied Language and Intercultural Studies, Dublin City

University, Republic of Ireland

Peter Reynolds, *CEO*, Translation Management Global, Poland

Florian Sachse, *Managing Director*, PASS Engineering GmbH, SDL International, Bonn,
Germany

Gerrit Sanders, XenoText, Herentals, Belgium

Prof. Dr. Klaus-Dirk Schmitz, Cologne University of Applied Sciences, Cologne, Germany

Prof. Dr. Uta M. Seewald-Heeg, Anhalt University, Köthen, Germany

Prof. Dr. Nitish Singh, Boeing Institute of International Business, John Cook School of Business,
Saint Louis University, St. Louis, Missouri, USA

Prof. Dr. Frieda Steurs, Lessius University College/K.U.Leuven University, Belgium

Thomas Vackier, *Localisation specialist*, Yamagata Europe, Gent, Belgium

Prof. Dr. Sue-Ellen Wright, Modern and Classical Language Studies, Kent State University,
Kent, USA

Dr. Jost Zetzsche, TM Marketplace, Sandpoint, USA

Dr. Andrzej Zydrón, *CTO*, XML-INTL, Bucks, UK

The *Journal of Internationalisation and Localisation* welcomes contributions on the following themes.

The main themes of JIAL include, but are not confined to:

Interoperability

Translation Quality Assurance

Development and compatibility of translation and localisation software with today's internationalisation and localisation's requirements

Web internationalisation

Translation marketing

Cultural adaptation

Development, implementation and certification of standards

Education, research and best practices in the areas of internationalisation and localisation

Interchange between course programmes and industry

Authoring

Software internationalisation

Localisation project management

Games localisation

Multimedia translation

The Open Source paradigm

Terminology management issues of internationalization and localization

JIAL is a peer-reviewed journal which is published once a year, and on the occasion of dedicated congresses.

For ordering print issues, please refer to <http://www.lessius.eu/jial/default.aspx>.

Acknowledgement

The Editor wishes to thank the Communication Department of Lessius University College, and especially Patricia Hellriegel who designed the web site for JIAL.

Preface

Research related with internationalisation and localisation is new and is more and more solicited by language businesses, software developers, translation agencies, international multilingual organisations, universities, language planning policy makers and standardisation institutes.

The *Journal of Internationalisation and Localisation* [JIAL] aims at establishing a worldwide discussion forum for both professionals and academics in the area of internationalisation and localisation. The scope of the journal is as broad as possible in order to target all the players in the internationalisation and localisation profession. The specific aim of the journal is to leverage the full range of information, from academic research results to the shop floor of today's language industries, and, conversely, to leverage business experiences in order to inform academic research.

This new journal's publication goals are driven by the scopes of the Localisation Industry Standards Association (LISA) and its education initiative, the LISA Education Initiative Taskforce (LEIT), offering the following opportunities and benefits to its expected audience:

Publicity and branding opportunities for GILT-oriented training programs;

Industry feedback enabling programs to remain abreast of industry trends;

Opportunities to disseminate critical information at both the academic and the industry levels;

A venue where educators, graduates, and industry can come together to foster burgeoning links between academia and industry.

JIAL aims not only to provide for standard academic publication practices, but also to adopt an editorial policy that is able to target academic research results efficiently towards

industry. Consequently, each new issue will appear in two publication modes: a pdf file format, accessible on line, and a print version on demand. The Editorial Board is committed to meet the needs for the presentation of academic content, while these more detailed papers will be complemented by the abstracts, and some key conclusions aimed at directly informing all partners involved in the internationalisation and localisation industry.

Hendrik J. Kockaert

Editor in Chief

The Journal of Internationalisation and Localisation

Volume I [2009]

ISSN (on line) 2032-6912

ISSN (print) 2032-6904

Edited by Hendrik J. Kockaert

Table of Contents

Mari D. González

Interactive Food & Beverage Marketing: Targeting Latino Youth in the Digital Age 4

Johann Roturier & Sabine Lehmann

How to treat GUI Options in IT Technical Texts for Authoring and Machine Translation 40

Miguel A. Jiménez-Crespo

The evaluation of pragmatic and functionalist aspects in localization: towards a holistic approach to Quality Assurance 60

Minako O'Hagan

Evolution of User-generated Translation: Fansubs, Translation Hacking and Crowdsourcing . 94

Barbara Inge Karsch

Profile of a Terminologist in Localization Environments..... 122

Sudhir H. Kale & Natalina Zlatevska

Globalization, Reterritorialization, and Marketing..... 150

Pablo Muñoz Sánchez

Video Game Localisation for Fans by Fans: The Case of Romhacking..... 168

Interactive Food & Beverage Marketing: Targeting Latino Youth in the Digital Age

Mari D. González

University of the Pacific

School of International Studies

Abstract

Recent interest in U.S.-based Latino/Hispanic populations and their cultures has led to fruitful research and increased attention. Comprehensive books on Hispanic marketing have been made available by multicultural marketing experts. Marketing that targets Latino/Hispanic youth has become a promising, specialized, and lucrative field. This study shows how cultural research has been utilized by marketers to target young Latinos/Hispanics, who are the most susceptible members of the ethnic group to new trends in technology. It investigates the type of cultural knowledge that marketing researchers are using to target Latino/Hispanic youth and the effectiveness of their interactive advertising campaigns. It also explores how the ever-growing access to digital media changes the way the food and beverage companies do business with Latino/Hispanic youth.

Mari D. González

The goal of this study is to combine academic and non-academic research on multicultural internet marketing in order to examine how food and beverage marketers have utilized cultural knowledge to reach Latinos/Hispanics. The study examines what this specific segment has to say about their online media habits and digital advertisement. Information on U.S. Latinos/Hispanics as online consumers seems to constantly change among marketers, and because of the constantly upgrading and renewing nature of technology, new findings will emerge to rebuke old ones. It is my endeavor to combine what researchers have studied on culturally relevant digital marketing with young Latino/Hispanic Americas' experiences and their preferences with internet, mobile phone, and food and beverage advertising.

A Note on Terminology

There is no current agreement on which of the two official terms “Latino”¹ or “Hispanic” is more accurate. Researchers (Davila, 2001; Oboler, 1995) indicate that either label “Hispanic” and “Latino” lump together and inaccurately define a group of people that span several generations, nationalities, and socioeconomic groups, and that “both terms are equally guilty of erasing differences while encompassing highly heterogeneous populations” (Davila, 2001, p. 15). For purposes of inclusion and based on the differences in meaning of each label among the people being categorized, I have chosen to use both labels and place them in that order for easy reading; however, the order is not to be read as a preference.

¹ “The term Latino was officially established in 1997 when the Office of Management and Budget published recommendations for classification of federal and ethnicity data indicating that the terms Hispanic and Latino were to be used interchangeably as of January 1, 2003” (Chong & Baez, 2005, p. 4).

Latinos/Hispanics in the U.S.

Latinos/Hispanics are the largest minority group in the United States. As of July 2007, they accounted for 15.1 percent of the total U.S. population (Wayne, 2009). It is projected that “the Latino labor force will increase by 77 percent during [2000 and 2020], while the non-Latino labor force will grow by only 9 percent” (Ramirez & de la Cruz, 2002, as cited in Chong & Baez, 2005, p. 6). Recent partnered advertising initiatives prove the extent of interest that these data hold for marketers. Major interactive media collaborations have shown increased attention being paid to Latino/Hispanic consumers as the new promising market. In April 2009, the online media news portal *Media Daily* (www.mediapost.com) published an article announcing that Terra USA, “the largest internet-access provider in the Americas,” (Valdes, 2002, p. 102) and a company that targets U.S. Latinos/Hispanics, had joined Interactive One, owned by network radio operator *Radio One* which mostly targets African-Americans to “cross-market and cross-sell their respective audiences” (O’Malley, 2009, ¶ 1).

The number of Latinos/Hispanics online has grown beyond initial expectations, surpassing the growth of the non-Latino/Hispanic market. In 2006, an online analysis (Williamson, 2006) projected that “US Hispanic internet users” would rise to 19.8 million by 2009. However, according to *comScore*, a global internet report (Kato, 2009), the number of online U.S. Latinos/Hispanics “climbed 50 percent faster than the overall Hispanic population in the past year” (¶ 1). By February 2009, it had reached 20.3 million, a six percent increase from the year before (Walsh, 2009, ¶ 1). “The Hispanic online population is now 11 percent of the total American market and in the last year [its growth] has significantly outpaced the rest of the market” (Wayne, 2009, ¶ 1).

Mari D. González

Online Latinos/Hispanics are younger than their mainstream counterparts (Williamson, 2006; Vence, 2005). Online research company *eMarketer* (Williamson) estimated in its 2005 report that the number of U.S. Latino/Hispanic internet users under the age of 35 would increase 33 percent from 9.1 million to 12.1 million in 2010, with one third under the age of 18. In 2008, the Latino/Hispanic online market was reported to be overall younger, with 56 percent between the ages of 18 – 24 compared to 34 percent of the general market in the same age range (Rios, 2008). Knowing these facts and figures makes it difficult to disagree with Assistant Professor of American Studies and Anthropology Davila's (2001) statement that, "Latinos are the hottest new market and those who target them will not regret it" (p. 1). Young online Latinos/Hispanics in the U.S. have a great market potential, and the food and beverage industry has stepped forward.

Drinks, Chips, Spanish-language Media, and American Football

One example of a recent marketing campaign by a food and beverage company occurred this year, when *PepsiCo* joined *Telemundo*, the number two ranked Spanish-language TV network owned by *NBC Universal*², to present two one-hour programs for Spanish-speaking audiences during *Telemundo*'s third annual Super Bowl special. The two programs, *Tazón Latino III* (Latino Super Bowl III) sponsored by *Doritos*, and a concert, *Pepsi Música Super Bowl Fiesta*, aired on Saturday, January 31, the day before Super Bowl. These events were heavily advertised online on *Telemundo*'s <http://msnlatino.telemundo.com> and *Univision*'s, the number one Spanish-language media network in the United States, owned by *Univision Communications, Inc.*³, <http://www.nflatino.com>, the online unit that operates the *NFL*'s Spanish-language website.

² 2 Hispanic Fact Pack, Annual Guide to Hispanic Marketing and Media, AverstisingAge, 2008 Edition.

³ 3 Hispanic Fact Pack, Annual Guide to Hispanic Marketing and Media, AverstisingAge, 2008 Edition.

Review of the Literature

Culture and Marketing

There is a greater-than-ever interest in Latino/Hispanic culture as a way to tap into their market potential “not just [because] the growing numbers.... [but because their commanding] buying power,” (Singh, Kwon & Pereira, 2003, p. 868) which is estimated to be at “just under \$800 billion,” for 2006 (Dodson & Kvicala, 2006, ¶ 1). In 2003, Latinos/Hispanics became the largest minority, surpassing African-Americans in number, resulting in a generation that is expected to influence the American economic, political and cultural landscape (Grow, Grover, Weintraub, Palmeri, Der Hovenesian & Eidman, 2004).

Research (Cui, 1997; Singh, Kwon & Pereira, 2003) establishes the relationship between ethnic segmentation and marketing advantage. It has been proven that understanding the consumers’ ethnic culture is the best way to communicate with and thus attract them (Cui, 1997; Korzenny & Korzenny 2005; Peñaloza & Gilly, 1999). Academic research (Cui, 1997; Peñaloza & Gilly, 1999; Singh, Baack, Pereira, & Baack, 2008; Singh, Toy & Wright, 2009; Torres & Gelb, 2002), internet analysts (Chang, 2009; Fadner, 2008; Medina, 2006; Vence, 2005; Williamson, 2006) and multicultural advertising professionals (Korzenny & Korzenny, 2005; Valdes, 2002; Vann, 2006) agree on the significance of “cultural relevance” when advertising to online Latino/Hispanic youth. A “culturally relevant” advertisement speaks to the collective values and/or ethnic identity of a particular cultural group.

PepsiCo’s collaboration with *Telemundo* demonstrated a strategy in line with current research for targeting Latinos/Hispanics in the U.S., which indicates that “the large growing U.S. Hispanic segment prefers culturally-adapted marketing messages” (Singh, Baack, Pereira &

Mari D. González

Baack, 2008, p. 225). Lisa Phillips, *eMarketer* senior analyst, explains that relevancy for Latinos/Hispanics is related to “understanding...their culture and traditions” (Fadner, 2008, ¶ 3).

Culture and the Web

Latinos/Hispanics in the U.S. share a common immigrant and acculturating experience (Vann, 2006), a common sociopolitical experience, and a common history of colonialism (Martin, 2005). They have a strong connection with their countries of origin and their loved ones back home (Vann, 2006). An effective website should therefore “account for differences in perception and attitudes between Hispanics and the [mainstream] market” (Vann, ¶ 9). A positive online experience ought to include an emotional connection to their heritage, which for example may explain Latin Americans’ “passion” for their national soccer teams (Vann).

Based on the 2006 *AOL Latino Cyberstudy*, Medina observes the differences in preferences in online behavior between acculturated and non-acculturated Latinos/Hispanics in the U.S. According to the report, “web habits” displayed by the acculturated group resemble U.S. white Americans’ or mainstream web habits; they visit entertainment, finance, and health-related sites, while the non-acculturated group prefer sites that inform them on Latin American events and on soccer (Medina).

Two Cultures

Palumbo and Teich (2005) recognize the relevance of a bi-dimensional acculturation model (Ryder et al., 2000, as cited in Palumbo & Teich), in which an ethnic population such as Latinos/Hispanics in the U.S. “retain[s] parts of [its] culture even as they gain American ideals and values” (p. 155). Ibarra (2009) observes, “Latinos live closer to their families’ countries of

origin and have access to Spanish-language television, radio, newspapers and magazines...[and] a universe of Web sites” (p. 84).

Researchers Palumbo and Teich (2005) suggest that marketers should adapt their messages in a way that represents aspects of both cultures —U.S. American and Latino/Hispanic— that form “[their] new Hispanic-American identity” (pp. 155-156). They conclude, “[a]dvertisements should include both American and Latino elements” (p. 156) and bilingual messages because this not only shows respect for and acknowledgment of their two cultural values, but it also shows their inclusion into the mainstream society (Palumbo & Teich). These suggestions could also apply to localization professionals.

Relevance of the Study

The relevance of this study is based on the premise that the number of second-generation Latinos/Hispanics is growing at a faster rate than that of first generation Latinos/Hispanics. Currently, there are more Latinos/Hispanics being born in the U.S. than the number of Latinos/Hispanics immigrating to the U.S. A study of digital advertising trends that target young U.S. Latinos/Hispanics should be useful to online marketers and localization professionals in further understanding a generation that has not been looked at as a separate segment. I focus on the food and beverage companies because this industry “[works] with a growing number of ad agencies, market research firms, and consulting groups that specialize in developing digital strategies for targeting [specific ethnic groups] in particular Latino... youth” (Chester & Montgomery, 2007, *Multicultural Marketing*, ¶ 1). In reflecting this need I pose the following question: How are food and beverage companies utilizing cultural knowledge to target Latino/Hispanic youth in the U.S. online?

Methods

In order to understand cultural perceptions and media use by young Latinos/Hispanics, this research relied on ethnographic interviews (Lindsley & Braithwaite, 1995). Lindsley and Braithwaite observe that ethnographic interviews “emphasize descriptive emic [meaningful to actor and culture-specific] information of people’s actual behaviors and their own interpretations of their communication interactions” (p. 202). As such, this study looks at interviews in which participants reported on their use of mass media, including internet and mobile phones, and their opinions on interactive and food and beverage advertising.

Ten bilingual college and high-school students each participated in a one-hour interview. They responded to an invitation by their Puente Project⁴ and English Language Development (ELD) advisors for an interview of media use by Latino/Hispanic youth. The face-to-face interviews were conducted between May 19 and May 26, 2009 in Pleasant Hill, Concord, and Fremont, California. Four of the participants were high school students and six were junior college students. There were two US-born participants; five from Mexico; two from Guatemala; and one from Bolivia with parents from Bolivia and Argentina. There were three males and seven females. Their ages ranged from 18 to 22.

Participants had to be students who identified as Latino/Hispanic and within the ages of 18 and 25. The written information provided to them was in English and Spanish including two consent forms and an interview questionnaire. They had the choice to answer the questions in English or Spanish. Three students requested to be interviewed in English and seven in Spanish. However, those who requested to be interviewed in Spanish had answers that included English-

⁴ A state-wide interdisciplinary junior college and high school program with counseling and mentoring components for underrepresented students in California. Retrieved from <http://www.puente.net/about/>

language words and phrases. The participants who requested English-language interviews had lived in the U.S. ten or more years. The interview consisted of 39 open-ended questions divided into seven sections: 1) TV, 2) Radio, 3) Mobile Phone Use, 4) Internet Use, 5) Social Network Sites (SNSs), 6) Internet Ads, and 7) Food and Beverage Advertising. Although this study focuses on digital media, the questionnaire included TV and radio to explore any connections between broadcast media and internet and mobile phone use.

Because technology and its utilization changes continuously, a new interview questionnaire was developed. The questionnaire was intended to demonstrate whether this small sample represented what multicultural advertising professionals have reported to date on Latino/Hispanic digital media usage and to lead the way for further research on bilingual Latino/Hispanic media experiences, internet advertisement, and localization strategies.

This study also gathers up-to-date information on multicultural digital media for Latinos/Hispanics in the U.S. Participants' answers were tabulated by looking for similarities, differences, and common themes. The challenges of this bilingual study were giving the participants the choice to answer in either or both languages and translating their answers into English. One certified interpreter whose first language was Spanish and one translator whose first language was English looked at the questionnaire translations to review the accuracy of translations. The translator suggested changes to the English-version questionnaire on back translation from translated Spanish-version to keep it "true" to the Spanish translation. She also looked at the English translations of Spanish quotes.

Results

The data were analyzed by tabulating participants' language of preference in different settings, number of hours spent on preferred media outlet, type of online activities, what participants considered to be relevant to their cultural identity, and their opinions on interactive and food and beverage ads.

Language of Preference: English, Spanish, Either One, or Spanglish⁵

Participants' language of preference was measured for different settings, and language options included: 1) Spanish; 2) English; 3) "either one," Spanish or English; and 4) Spanglish. All of the participants preferred speaking Spanish at home; however with friends, four preferred Spanish, three Spanglish, two "either one," and one English.

At school, six respondents preferred using English, two "either one," one Spanish, and one Spanglish. Five of respondents read the news in English, three in Spanish and two in "either one." Five of respondents browsed the internet in English, and five in "either one." Out of the eight participants who reported being employed, five stated that at work they spoke "mostly" English, two "mostly" Spanish, and one "either one." At work participants did not choose one language over another. They explained that they were required to speak the preferred language of their supervisors, coworkers, and customers.

TV

Interview participants watched an average of three hours of TV a day. Their selection of programs included 1) Spanish-language programs from their countries of origin (for example, *Canal Sur*, a South American cable channel or Mexican movies from the 1950's golden era); 2)

⁵ "The blend between Spanish and English found in Hispanic or Latino communities in the United States." Ardila, A. (2005). Spanglish: An anglicized Spanish dialect. *Hispanic Journal of Behavioral Sciences*, 27(1), 60-81. Retrieved from <http://hjb.sagepub.com/cgi/content/>

Spanish-language programs (*telenovelas*⁶, preferred by the females; and soccer games, preferred by the males) on U.S. media outlets *Univision*, and on *Telemundo*; 3) bilingual video and music-centered programs such as *MTV Tr3s*, a bicultural entertainment network owned by *MTV Networks* (*Viacom*); 4) English-language documentary and “kids” TV channels (*Discovery*, *Disney*, *Nickelodeon*) and 5) Latino/Hispanic-themed programs on mainstream media (reruns of the George Lopez TV series).

Radio

Participants reported listening to radio in their cars and/or online an average of 1.3 hours per day. Their music preferences fell into three main categories, 1) Pop in English (rock, alternative rock, hip hop, urban), 2) Latin (*reggaeton*, *bachata*, *rock en español*), and 3) regional from their country of origin (*banda*, *rancheras*, *folklórico*, *marimba*).

The most frequently mentioned local radio station was San Francisco Bay Area’s *La Kalle*, owned by *Univision Communications, Inc.*, which appeals to young bilingual urban Latinos/Hispanics. Two participants described *La Kalle* as bicultural and modern, and said that they listen to it because 1) “ponen música nueva, del género urbano [they play new music of the urban genre];” 2) “[porque me entero] de las noticias, anuncios, eventos culturales para Latinos [como] ‘Cinco de Mayo,’ [y] ‘Día de las Madres’ [(because I find out about) news, cultural events for Latinos (such as) ‘Cinco de Mayo,’ (and) ‘Mother’s Day’].” One participant reported that she rarely listened to music on the radio. She downloaded her favorite music and burned it on CDs. She commented, “Most of the music I listen to is from the internet; mostly in English, *reggaeton*, pop, and hip pop. I download the music that I listen to on TV video programs [like]

⁶ A limited-run television serial melodrama. Retrieved from <http://en.wikipedia.org/wiki/Telenovela>

Mari D. González

MTV Latino [MTV America Latina or] ‘*Dos mundos*’ [mun2]⁷.” She added, “Todo el mundo las escucha [como] mis amigas [Everybody listens to them (like) my girlfriends].”

Mobile Phone Use

Out of the ten participants, eight had access to the internet on their mobile phones. One respondent did not own a cellular phone and another did not pay to have internet access. Of those with access to mobile internet, half reported using it for entertainment purposes: 1) “[I use] email, IM, música [music downloads], [and] sistema de navegación [GPS navigation system];” 2) “[I use] *hotmail messenger*, [para] bajar música de *FunForMobile* [porque tiene] la música que me gusta gratis” [(to) download music from *FunForMobile* (because it has) the music I like free];” 3) “[I use] *Twitter*, I recently opened [an account]. I follow a lot of artists [for instance] *Panda* from Monterrey. [They play] rock en español;” 4) “I use internet on my cell phone [to] download music, [use] email, *MySpace*, and *Hotmail*.” The other four who paid to access internet on their cell phones reported using it mainly for emergencies such as GPS navigation system or to check their email when away from home.

When asked if they received advertising via their cell phones, five out of the eight who used cell phones, said “yes,” and four reported receiving advertising only from their wireless phone providers including *Verizon*, *T-Mobile*, *Metro PCS*, *AT&T*. One of the participant who reported having received ads from non-wireless phone companies stated, “*Domino’s Pizza* me envía promoción, ‘el especial de la semana.’ [Creo que obtuvieron mi número cuando] una vez ordené pizza por [mi] teléfono [móvil]. [*Domino’s Pizza* sends me promotions, ‘the special of

⁷ The first national cable network to offer bilingual programming for young U.S. Latinos. Retrieved from <http://holamun2.com/press/pepsi-musica-launch>

the week.’ (I believe they got my number when) once, I ordered pizza from (my mobile) phone].”

When asked if they would mind receiving ads on their cell phones, six out of the eight who used a cell phone said they “would not mind.” Three stated that as long as the ads were about products they like it was ok, 1) “No importa si es otra compañía pues me doy cuenta; si es algo que me interesa [I don’t mind if it’s another company that way I find out if it’s something I’m interested in];” 2) “[It] depends on what the ad is. Something I like, no problem;” 3) “No me molestaría recibir ‘texts’ de productos y eventos que me gustan [I wouldn’t mind receiving texts on products and events that I like].” One participant clarified that he would not like to receive ads from the same company on a daily basis, “No me molestan [pero] si me lo mandaran todos los días lo bloquearía [They don’t annoy me (but) if they were to send them everyday, I’d block it].”

Of the cell phone users seven reported having sent and/or forwarded text messages to their friends about concerts, promotions, new information from the internet, school-related events, inspirational quotes, party invitations, and/or ads on missing children. The promotion they reported receiving and forwarding via text message most frequently was for “a free breakfast at *Denny’s*.” One participant explained, “*Denny’s* tenía desayunos gratis y ‘textie’ a [mis] conocidos. Me enviaron un text en español; fue una combinación [de promoción] de TV en español. Fueron al desayuno como cinco amigos [*Denny’s* had free breakfasts and (I) texted (my) acquaintances. A text was sent to me in Spanish; it was a (cross) promotion with Spanish TV. About five of my friends went to the breakfast].” Another participant indicated, “Siempre

Mari D. González

les digo a mis amigos. Los Latinos somos muy comunicativos [I always let my friends know.

We Latinos are very communicative people].”

Internet

Interview participants reported spending an average of four hours a day online. Their top activity was interacting on social network sites (SNSs). The second most popular activity was “music downloading,” followed by reading and/or watching the news and sports (soccer and “*lucha libre*,” [wrestling]) news, email, and chatting. Other online activities mentioned included shopping, using search engines, retrieving driving directions, reading Spanish-language media (Univision.com), and soccer news. Respondents also reported visiting *Univision*’s online *telenovela* forums. *Limewire* was mentioned by three students as their preferred site to download music. Other music download sites mentioned were *Ares*, *Videorola*, a regional Mexican online radio and *Sideral*, a Guatemalan online radio, and *iTunes*.

Social Network Sites (SNSs)

Only one participant reported not having an active profile on any social network sites. Eight reported having an active profile on *MySpace*; three on *Facebook*; two on *Hi5*; and one on *MetroFlog*, a Spanish-language photolog website. Among the reasons mentioned in favor of *MySpace* were, 1) “Es el más popular, todos mis amigos lo tienen [It’s the most popular, all my friends have it (including) the ones from Mexico];” 2) “[Tiene] más aplicaciones, oportunidades para expresarte, decorar, fotos [It has more applications, opportunities to express yourself, decorate, pictures];” 3) “Everybody has *MySpace*, y se la pasan allí [they spend their time there]. If my friends don’t answer the phone, they’re on *MySpace*;” 4) “[It’s] easier to use, manage [and] put pictures, [I] get a lot of help, [answers to my] questions, editable pages, it’s all decorated,

different, music, pictures, [information] about me;” 5) “Puedes poner fotos las que quieras, escuchar música, escoger colores, [hacerlo] llamativo [You can post pictures all you want, listen to music, choose colors, (make it) attractive];” 6) “*MySpace* en español para comunicarme con amigos en los dos idiomas [*MySpace* in Spanish to communicate with friends in both languages];” 7) “I like *MySpace* more because I like to follow bandas [bands] and [I] get to hear [their] music;” 8) “Puedes crear la pagina a tu manera. Me da la oportunidad de expresarme a mi manera. Lo tengo extravagante [You can create the page your way. It gives me the opportunity to express myself my way. I have it extravagant].”

Internet Ads

Interview participants reported recalling online ads for recently-released movies, fast food products, weight loss products, skin care (acne) products, concerts, dating sites and free laptops. The most frequently recalled ad was for weight loss products, mentioned by three female respondents. Two respondents reported being annoyed by internet ads, saying that 1) “Me molestan porque estoy [they bother me because I’m] ‘chatting’ [and] they’re in my way;” and 2) “Once [I clicked on it], me bajo muchos [it got me a lot of] viruses.” Three respondents reported downloading wallpaper related to the product’s promotion.

Interactive Food and Beverage Ads

Three respondents favored *Coca Cola*’s ads. Two describe them as creative: 1) “Destapas la botella y te dirige a su sitio... tiene mucha música de animación, alegre [You open the bottle and it leads you to its (web) site... it has a lot of animated music, cheering];” 2) “I like Coke [ads. They] bring your senses to the product.” Another described *Coca Cola*’s promotions in Latin America as “especial [special]” and “stronger” than in the U.S. “[*Coca Cola*] traía

Mari D. González

grupos de rock y ponían conciertos en las [*Coca Cola* used to bring rock bands and had concerts in the] high schools.”

Two respondents mentioned *McDonald's* ads as their favorite online ad. One noted the importance for companies to give back to the Latino/Hispanic community, “[Mi anuncio favorito es] sobre una joven [estudiante] que ganó una ‘scholarship’ [y dice] ‘Gracias a *McDonald's* pude ir a la escuela.’ Lo más relevante es que ayuden a la comunidad [Latina/Hispana]. Las compañías ganan mucho de nuestra comunidad, si no [nos] van a ayudar... para que consumir. [My favorite ad is about a young student who won a scholarship and says, ‘Thanks to *McDonald's* I was able to go to [college].’ The most relevant is [for companies] to help the Latino/Hispanic community.” Another respondent liked the fact that *McDonald's* has the same ads in Spanish and English. However, he explained, “Me identifico más con el de español [I identify more with the Spanish one].”

One respondent described a *Redbull* energy drink ad, “[It’s] funny, gracioso [funny], cartoons. [Muestra] al abuelito bailando en la mesa y el nieto dice, ‘mejor no le hubieramos dado la bebida’ [(It shows) the grandfather dancing on the table and the grandson says, ‘We shouldn’t had given him the drink’].”

Online Promotions

Five participants stated being registered on promotional websites. Two mentioned the *Coca Cola* website (www.mycoke.com) where members can earn promotional points. They reported that downloading free music was their main reason for registering. Another respondent stated that he was registered on the *Domino's Pizza's* website “para obtener cupones [to get

coupons].” The other two mentioned *Victoria’s Secret* website, as requested by her mom and *Mary Kay* website for the informative lessons on make up.

Five respondents stated that they were not registered on any promotional website. Two of them noted the sites’ requirements as the reason for not registering: 1) “there are a lot of requirements;” 2) “too many requirements.” One expressed her disbelief to win something, “I’ve never done it. I don’t believe I have a good chance [to win].”

Six respondents reported that they did not have any language preference for internet ads. Three participants reported preferring online ads in Spanish: 1) “Spanish [ads] are funnier. I understand humor in Spanish.” 2) “[I] prefer [them] in Spanish so that my parents can understand; to persuade my family.” 3) “Spanish, to understand the offer... humor in English is not as funny. [En español] le ponen el vocabulario mas figurative [(In Spanish) they use a more figurative vocabulary].” Two reported preferring bilingual ads and one recognized that they were uncommon: 1) “Es mejor ‘mixteado’ porque resalta el español, se nota. Puro español no hay palabra chusca que resalte [It’s better mixed because Spanish stands out, it can be noticed. (In) pure Spanish there isn’t any funny word that stands out].” 2) “Hay pocos. No hay muchos. Hay mas en ingles [There are few. There aren’t many. There are more in English].

Cultural Relevance in Advertising

Interview participants mentioned several “culturally relevant” online elements, which included Latin music, brighter colors, inclusiveness, cultural values, bicultural Latino/Hispanic role models, and successful Latino/Hispanic individuals in mainstream media. Their comments on these culturally relevant elements included the following:

- Latin music makes websites “Más llamativa [More inviting].”

Mari D. González

- Inclusiveness: “[I] hate discriminatory commercials that [show] that only Mexicans speak Spanish [or] that only Mexicans eat tacos.” “[Me gustan] películas donde ponen música en español. Nos hace sentir bien porque si nos toman en cuenta [(I like) movies where they put music in Spanish. It makes us feel good because they take us into account].”
- Cultural values: “It’s important to show the Latin flavor... for Spanish-speakers. It makes me feel proud. Language does not make [as much] difference but values, [the] importance of family, [being] hardworking, being a good friend.”
- Bicultural Latino/Hispanic role models: “Me gusta [George Lopez] porque muestra una familia Hispana ‘mixteada’ con como es la cultura de Estados Unidos.” [I like (George Lopez) because it shows a Hispanic family mixed with the way the U.S. culture is].”
“Enrique Iglesias, su música sigue siendo Latina pero con pop. Cuando graba discos siempre incorpora la versión en español. No se aleja completamente de su cultura [Enrique Iglesias, his music is still Latin but with pop. When he records (CDs) he always incorporates the Spanish version. He doesn’t stray away completely from his culture].”
“Oscar de la Hoya, siempre ayuda a los Latinos, es importante... que uno no se olvide de sus raíces y que ayude a su gente [Oscar de la Hoya always helps Latinos, it’s important... that one doesn’t forget his roots and helps his own people].”
- Successful Latino/Hispanic on mainstream media, “I wish there were more Latinos in Hollywood, we’d get more respect [like] Gael Garcia Bernal. I was surprised to see his latest movie on a mainstream theater. It gives him exposure. It says a lot.”

Discussion

Research points out three main cultural values when targeting online U.S.

Latinos/Hispanics: 1) preference for bilingual content; 2) *familismo*; and 3) the use of celebrities as spokespersons. While multicultural internet marketers also recommend bilingual websites, they note two more recent trends: 1) the extensive use of mobile internet; and 2) the widespread use of entertainment and social network sites by young Latinos/Hispanics.

Bilingual Websites Preferred by Bilingual Youth

The majority of online Latinos/Hispanics prefer English content. Yet, bilingual U.S.-born or second generation and 1.5-generation Latinos/Hispanics, who immigrated as children, prefer English websites that include Spanish because those sites speak to their cultural identity and make them feel included (Vann, 2006). In fact, researchers (Singh, Baack, Kundu, & Hurtado, 2008) argue that “[l]anguage in this case... Spanish, tends to be the most visible manifestation of U.S. Hispanic identity” (p.2). According to Williamson (2006), Latinos/Hispanics appreciate “quality Spanish-language content online.... [It resonates with their] cultural pride and a feeling of community” (p. 17).

Multicultural marketing analysts (Singh, Baack, Pereira & Baack, 2008; Vann, 2006; Wasserman, 2009; Williamson, 2006;) advise advertisers to include both Spanish and English to reach the whole spectrum of Latino/Hispanic online users. Most U.S. Latino/Hispanic youth are comfortable with either English or Spanish advertisements and marketing professionals understand that simply translating or adapting websites does not work for Latinos/Hispanics (Korzenny & Korzenny, 2005; Valdes, 2002; Vann 2006) because they are fully bilingual and bicultural (Valdes).

Mari D. González

Familismo: the Maximum Representation of Latin/Hispanic Collectivism

Young Latino/Hispanic Americans hold onto the *familismo*, the strong value Latinos/Hispanics place in family orientation. They acknowledge being swayed by parents and grandparents to watch a particular TV program or registering for promotional websites of products their parents like. They also appreciate TV programs and ads that allude to *familismo*. *Familismo* and collectivism play an important role and are central to their culture and traditions (Chong & Baez, 2005; Davila, 2001; Korzenny & Korzenny, 2005; Singh, Kwon & Pereira; 2003, Valdes, 2002; Vann, 2006). According to Chong and Baez (2005), familismo “is one of the most important Latino cultural values” (p. 39). Online advertising professional Lee Vann (2006) advises marketers to “craft an online experience that appeals to strong family principles.... [for the site to ‘feel] Hispanic” (§ 4). Companies that reach out to minorities capitalize on peer and family pressure because collectivist societies tend to share more information and depend more on word-of-mouth referrals (Singh, Kwon & Pereira, 2003).

Use of Celebrities as Spokespersons to sell Food and Beverages

Research (Singh, Baack, Pereira & Baack, 2008; Torres & Gelb, 2002) recommends the use of audience-like spokespersons or models to increase ad receptivity. Torres and Gelb, suggest “accommodation” or the use of people “similar to target audience to increase the likelihood of reaching communication objectives” (p. 71).

Among the top 50 food and beverage companies that advertised in Hispanic media in 2008⁸ as measured by U.S. media spending included four food and beverage companies: *McDonald’s* (No.6); *SABMiller* (No. 24); *PepsiCo* (No. 27); and *Coca Cola, Co.*, (No. 43).

⁸ Hispanic Fact Pack, Annual Guide to Hispanic Marketing and Media, AvertisingAge, 2008 Edition.

Food and beverage companies effectively utilize “culturally relevant” advertising to target Latino/Hispanic youth. They employ Spanish-language media, well-known Latino/Hispanic celebrities, and *telenovela*-inspired ads to promote their products. The list of companies includes *Carl’s Jr.*, *Coca Cola*, *Tecate* (*Heineken* Mexican beer), *McDonalds*, and *Pepsi*. Trade magazine, *Mediaweek* (Bachman, 2009) announced that *Carl’s Jr.* was launching a TV campaign for affiliates at *Univision* and *Telefuturo*, a Spanish-language broadcast network owned by *Univision Communications, Inc.*, from March 23 through June 23, 2009. Benchman (2009) describes *Carl’s Jr.* ad “disguised as mini-novela” (§ 1). *Carl’s Jr.* media company approached *Univision’s* “*telenovela* experts” for the ad production. The *telenovela* characteristics include, 1) the characters, “a scantily clad young woman... [and] a ruggedly handsome man;” 2) the suspense, “[being] interrupted by the woman’s former lover;” and 3) concludes with a happy ending, with the woman enjoying the burger with two men” (§ 4).

In March 2009, *Coca Cola* began airing a Latino/Hispanic-targeted campaign with a *telenovela*-related promotion. This is not the first time *Coca Cola* joined a Spanish-language *telenovela* fan club for cross-promotion, but the fifth. For one week, the promotion participants entered sweepstakes to win *telenovela* memorabilia and a chance to meet *telenovela* stars in person (Parpis, 2009).

Pepsi relies on promotional music concerts. According to de Lafuente (2008), “*Pepsi*... has made a significant push in music programming,” (§ 5) to target young bilingual Latinos/Hispanics with *PepsiMusica*, named after its main sponsor, is a bilingual entertainment program that showcases top Latin pop music performers from Latin America and the U.S. aired on *Telemundo’s mun2* network.

Mari D. González

Tecate, Heineken's Mexican beer, targets Latinos/Hispanics with dual campaigns and separate English and Spanish ads. While the Spanish-language ads were aired on *Univision*, *Telefutura*, and *Telemundo* affiliates, the English-language ads were aired on *ABC*, *CBS*, *CW*, *Fox*, and *NBC*. The Spanish-language ads are “aimed at recent immigrants from Mexico.... [and] the English-language ads celebrate the influence of Mexican-Americans in the U.S. culture” (Greenberg, 2009, ¶ 1). The later campaign features boxing legend Oscar de la Hoya on one of the spots (Greenberg).

Bicultural Campaign

In 2007, *McDonald's* designed a campaign for bilingual Latinos/Hispanics called “Flying Fry” which features two young Latino/Hispanic actors. *McDonald's* ads did not utilize celebrities or *telenovela* formats. Instead, it ran an “elusive and breakthrough” bicultural campaign that used TV, radio, and online ads to “[connect] with the growing acculturated segment of Hispanics, especially second generation” (de Lafuente, 2007, ¶ 5). The campaign centered on Latinos/Hispanics between the ages of 16 and 19. They combined Latino/Hispanic celebrities, Latin music, and Spanish-language with a mainstream message. The ads aired on English-language networks including *Comedy Central*, *Fox*, *MTV* and *VH1*. The interactive online banners came up during urban-type programming. Tommy Thompson, the ad agency's president, observes that bilingual Latinos/Hispanics do not connect with mainstream English-language advertisement because, “they don't find the message relevant” (de Lafuente, 2007, ¶ 7).

AdvertisingAge's Hispanic Marketing and Media guide for 2008 reported that *McDonald's* Corp., spent \$89,657.6; *PepsiCo.*, spent \$41,054.7; and *Coca Cola Co.*, spent \$26,990.8 thousands of dollars in U.S. “Hispanic” media for 2007.

Current Trends on Online Advertising

Mobile Internet and Entertainment Media

Latino/Hispanic youth enjoy having access to Internet on their cell phones and they use it widely for entertainment purposes. Media analysts (Chang, 2009; de Lafuente, 2008; Kutchera 2009; Loechner, 2008) note that Latinos/Hispanics are the highest users of mobile media. Chang observes “U.S. Hispanics are more engaged with their mobile phones than Americans overall” (¶ 6). In 2009 *comScore M:Metrics* reported (Kutchera) that 71 percent of Latinos/Hispanics consume mobile online content while the “market-average” consumption is 48 percent (¶ 2). De de Lafuente observes, “part of this boils down to demographics: 51.5 percent of U.S. Hispanics with phones are between 18 and 34 years old while 31 percent of the general ‘mobile’ population falls within that range” (¶ 3). According to BIGresearch’s 2008 report (Loechner, 2008), Latinos/Hispanics topped the list of regular cell phone users of “the form of new media used the most...with 57%... [compared to] 53% of African Americans, 53.9% of Asians and 49.4% of Whites” (¶ 60).

Loechner (2008) argues that “minorities.... are more likely to use iPods, text on cell phones, play video games, use video/picture phones, instant messaging online and watch videos on cell phones... than Whites” (¶ 7). Reports (de Lafuente, 2008) on Latinos/Hispanics as “more sophisticated mobile users” drove the telecommunications company Sprint to launch a new cell phone that includes “GPS navigation, music download capabilities, text messaging, web browsing, live TV, picture taking/sending and voice activation” (¶ 7) and was set to be heavily campaigned to attract Latinos/Hispanics (de Lafuente). Latinos/Hispanics do not mind receiving

Mari D. González

mobile ads even pass them to their friends. They “tend to notice and respond well to ads on cell phones” (Kutchera, 2009, ¶ 2).

Online Entertainment Media and Social Network Sites

Young Latinos/Hispanics tend to socialize online more through social network sites and instant messaging sites than non-Latinos/Hispanics. *MySpace* and *MySpace Latino* executive Manny Miravete observes, “Social media... [is] the fabric of Hispanics online” (O’Leary, 2009, ¶ 3). Miravete reported that 65 percent of online Latinos/Hispanics use social network sites (O’Leary, 2008).

Williamson (2006) argues that Latinos/Hispanics spend more time chatting and/or using instant messaging than the general population online. Research firm, Mintel reports (Greenberg, 2009) “Hispanics are more likely to have profiles on social network sites than non-Hispanics... [and] it is young Hispanic consumers who are driving the trend” (¶ 4). Market research company Mintel (Greenberg, 2009) reports that “Hispanics spend nearly twice as much time -- 5 hours -- using networking websites or chatting via instant messaging than any other demographic group” (Greenberg, 2009, ¶ 2).

Conclusion

Based on the reported findings, online advertisers, internet marketers, and localization professionals wishing to target bilingual and acculturated Latinos/Hispanics in the U.S. ought to account for their “hybrid” cultural identity which is formed by American and Latino/Hispanic values. Localization professionals and online marketers can successfully reach this demographic group by including content that uses a blend of mainstream American and Latino/Hispanic cultural elements. For an online ad, an internet marketing campaign or a website to be

“culturally relevant” to this sub-group, it needs to include ethnic identifiers that are immediately recognized and collectively accepted by its members as Latino/Hispanic such as, Spanish-language words or phrases within English content; popular modern Latin/Hispanic music; Latino/Hispanic spokespersons, celebrities or “audience-like” models; and images that portray a Latino/Hispanic family.

Including these “culturally relevant” elements on ads or websites is more effective than Spanish-language only or simply translated content when targeting bilingual U. S. Latinos/Hispanics. Having the same ad or web page in both languages, English and Spanish, is not necessary for this demographic. It can be redundant. They are fluent in English and prefer to navigate the internet in English. Yet, a mainstream ad or website in English-language only does not resonate with them as much as a bilingual or Spanglish one because it includes the two languages they regularly interact with. Adding commonly-used Spanish-language words or phrases to an ad or web content is more attractive for bilingual and bicultural Latinos/Hispanics. It makes them feel included into the mainstream American society.

Popular bicultural Latino/Hispanic-American celebrities are effective spokespersons for this demographic on marketing campaigns because they are seen as role models. These celebrities are immersed in the American mainstream culture and at the same time they display a level of attachment to Latino/Hispanic values which is comparable to this sub-segment of bilingual young Latinos/Hispanics. Bilingual and bicultural second-generation Latinos/Hispanics are in tune with both cultures and as consumers they are best reached with online marketing and digital campaigns that are bilingual and bicultural as well. The best examples of this model are music-themed websites and TV video programming channels on

Mari D. González

which their hosts interview performers in English but incorporate some words and phrases in Spanish and the performers sing in both languages too.

Internet marketers and localization professionals can highlight *familismo* on an English-language website to give it a Latino/Hispanic feeling with pictures of Latino/Hispanic families that include grandparents and other extended family members.

Other online widespread elements recognized by Latino/Hispanic in the U.S. are brighter colors and on social network and interactive sites, the option for artistic and self-expression. Because of second-generation Latinos/Hispanics' affinity with soccer and *telenovelas*, a *telenovela*-inspired and a soccer-themed commercial or marketing campaign will strongly resonate with of Latinos/Hispanics in the U.S. To increase the effectiveness when advertising, promoting, and selling to this sub-segment, it is recommended opting for brighter colors and including applications for online interaction such as forums and social network sites and alternatives for the user to customize and creatively personalize her individual page.

The growing number of second-generation bilingual Latinos/Hispanics in the U.S. uses the latest digital technology such as mobile internet, online media applications and cell phone text messaging and makes extensive use of mobile media and social network sites. This market segment welcomes cell phone text ads as long as those are on products and services relevant to their age and lifestyles. Online advertisers and internet marketers that target them can effectively utilize cross-promotions that include TV, internet and mobile phone. Offering free-music downloads greatly increases the campaign's success given that this demographic makes the most of their mobile phone by downloading their favorite songs.

References

- Bachman, (2009, March 23). Carl's Jr. sizzles with Univision novela pitch. *Brandweek.com*. Retrieved from http://www.brandweek.com/bw/content_display/news-and-features/hispanic-marketing/e3i710f01ed18601fd49b46521626dc6b1b
- Chang, R. (2009, January 26). Mobile marketers target receptive Hispanic audience. *AdvertisingAge*, 18.
- Chester, J., & Montgomery, K. (2007). *Interactive food & beverage marketing: Targeting children and youth in the digital age*. Berkeley CA: Berkeley Media Studies Group.
- Chong, N. & Baez, F. (2005). *Latino culture: A dynamic force in the changing American workforce*. Yarmouth, ME: Intercultural Press.
- Cui, G. (1997). Marketing strategies in a multicultural ethnic environment. *Journal of Marketing Theory & Practice*, 5(1), 122-134.
- Dávila, A. (2001). *Latinos, Inc.: The marketing and making of a people*. Berkeley, CA: Berkeley University Press.
- de Lafuente, D. (2007, December 1). *McDonald's tempts Latinos with 'flying fry.'* *Brandweek*. Retrieved from http://www.brandweek.com/bw/eseach/article_display.jsp?vnu_content_id=1003688127
- de Lafuente, D. (2008, April 4) Report: McD's, Pepsi score best with young Hispanics. *Brandweek*, 49(16), 8-81/2. Retrieved from EBSCO database.
- de Lafuente, D. (2008, April 7). Marketers are taking advantage of the fact that Hispanics are ahead of the general market on mobile date use. *Adweek.com*. Retrieved from http://www.adweek.com/aw/content_display/special-reports/other-reports/

Mari D. González

e3i7d6885cb31509ec4623a2dddb22f9d7d

Dodson, D. & Kvicala, J. (2006, September 1). Hispanics will top all U.S. minority groups for purchasing power by 2007, according to annual buying power study from UGA's Selig Center for Economic Growth. *News at Terry College* 1(1). Retrieved from http://www.terry.uga.edu/news/releases/2006/buying_power_study.html

Fadner, R. (2008, December 18). Online marketers target minorities. *MediaPost Publications*. Retrieved from http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=96964

Greenberg, K. (2009, March 25). Hispanics' web savvy surpassing others. *MediaPost Publications*. Retrieved from http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=102873

Greenberg, K. (2009, May 2008). Heineken touts Tecate Light in dual effort. *MediaPost Publications*. Retrieved from http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=105736

Grow, B., Grover, R., Weintraub, A., Palmeri, C., Der Hovenesian, M., & Ediman. (2004, March 15). Hispanic nation. *Business Week*, 58-70. Retrieved from EBSCO database.

Hernandez, S. A. & Newman, L. M. (1992). Choice of English vs. Spanish language in advertising to Hispanics. *Journal of Current Issues and Research in Advertising*, 14(2), 35-45.

Kato, D. (2009, April 18). Retail 3.0: Hispanic online audience grows. *MyLatinoNews.com*. Retrieved from <http://mylatinonews.com/2009/04/retail-30-hispanic-online-audience-grows/>

- Korgaonkar, P., Silverblatt, R., & Becerra, E. (2004). Hispanics and patronage preferences for shopping from the internet. *Journal of Computer-Mediated Communication*, 9(3), 00-00.
- Korgaonkar, P. , Silverblatt, R. , & O'Leary, B. (2001). Web Advertising and Hispanics. *Journal of Consumer Marketing*, 18(2), 134-152.
- Korzenny, F. & Korzenny B.A. (2005). *Hispanic Marketing: A cultural perspective*. Burlington, MA: Elsevier Butterworth-Heinemann.
- Kutchera, (2009, January 15). Connecting via mobile: Speak their 'edioma.' Retrieved from http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=98416
- Lindsley, S. L., & Braithwaite, C. A. (1996). "You should 'wear a mask': Facework norms ins cultural and intercultural conflict in maquiladoras. *International Journal of Intercultural Relations*, 20(2), 199-225.
- Loechner, (2008, May 28). Minorities dominate use of new media. *MediaPost Publications*. Retrieved from http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=83414
- Martín L.A. (2005). *Latino vs. Hispanic: The politics of ethnic names*. Thousand Oaks, CA: Sage Publications.
- Medina, H. (2006, September 18). Online Hispanic's web habits vary by acculturation. *HispanicBusiness.com*. Retrieved from http://www.hispanicbusiness.com/news/2006/9/18/online_hispanics.htm
- Ocaña, (2009, June/July). Doing it our way: The way Latinos assimilate could be changing the very basis of what it means to be American. *Latina*, 82-84.

Mari D. González

Oboler, S. C. (1995). *Ethnic labels, Latino lives: Identity and the politics of (Re)presentation the United States*. Minneapolis, MN: University of Minnesota Press.

O'Leary, N. (2009, April 6). Why social media is the 'fabric' of Hispanics online.

Brandweek.com. Retrieved from http://www.brandweek.com/bw/content_display/news-and-features/hispanic-marketing/e3i51215d72cd029dc275e27e2719bf4555

O'Malley, G. (2009, March 19). Hispanics trail other demo groups online. *MediaPost*

Publications. Retrieved from http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=102475

O'Malley, G. (2009, April 13). Terra, interactive one partner for Hispanic social network.

MediaPost Publications. Retrieved from http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=104013

Palumbo, F., & Teich, I. (2005). Segmenting the U.S. Hispanic market based on level of acculturation. *Journal of Promotion Management*, 12(1), 151-173.

Parpis, (2009, March 31). Coke tells Hispanics: 'Unleash your dreams.' *Brandweek.com*.

Retrieved from http://www.brandweek.com/bw/content_display/esearch/e3i8e9e151d61b41dcaf875a6b709a9710b

Peñaloza, L., & Gilly, M. (1999). Marketer acculturation: The changer and the changed. *Journal of Marketing*, 63(3), 84-104.

Rios, F. (2008, August 7). Understanding and engaging Hispanic consumers. *MediaPost*

Publications. Retrieved from http://www.mediapost.com/publications/index.cfm?fa=Articles.showArticle&art_aid=88080

- Singh, N., Baack, D., Pereira, A., & Baack, D. (2008). Culturally customizing websites for U.S. Hispanic online consumers. *Journal of Advertising Research*, 48(2), 224-234.
- Singh, N., Baack, D. W., Kundu, S. K., & Hurtado, C. (2008). U.S. Hispanic consumer e-commerce preferences: Expectations and attitudes toward web content. *Journal of Electronic Commerce Research*, 9(2), 162-175.
- Singh, N., Kwon, I., & Pereira, A. (2003). Cross-cultural consumer socialization: An exploratory study of socialization. *Journal of Psychology & Marketing*, 20(10), 867-881.
- Singh, N., Toy, D., & Wright, L. (2009). A Diagnostic Framework for Measuring Web-site Localization. *Thunderbird International Business Review*, 51(3), 281-295.
- Torres, I. M., & Gelb, B. D. (2002, December). Hispanic-targeted advertising: More sales?. *Journal of Advertising Research*, 69-75.
- Ueltschy, L. C., & Krampf, R. F. (1997). The influence of acculturation on advertising effectiveness to the Hispanic market. *Journal of Applied Business*, 13(2), 87-102.
- Vann, L. (2006, June 23). An overview of the Hispanic online market. *Chief Marketer*. Retrieved from http://chiefmarketer.com/disciplines/online/Hispanic_online_market_06232006/
- Vann, L. (2006, July 10). Hispanic online best practices: In-language. *Chief Marketer*. Retrieved from http://chiefmarketer.com/disciplines/multicultural/Hispanic_online_in-language_07102006/
- Vann, L. (2006, July 21). Hispanic online best practices: In culture. *Chief Marketer*. Retrieved from http://chiefmarketer.com/disciplines/multicultural/Hispanic_in-culture_07212006/
- Valdes, M. I. (2002). *Marketing to American Latinos: A guide to the in-culture approach* (Part Two). Ithaca, NY: Paramount Market Publishing, Inc.

Mari D. González

Vence, D. (2005). Net Serves as Best Tool to Connect with Hispanics. *Marketing News*, 39(14), 29-37.

Walsh, M. (2009, April 16). U.S. Hispanic online audience surpasses 20 million. *MediaPost Publications*. Retrieved from http://www.mediapost.com/publications/?fa=Articles.showArticle&art_aid=104268

Wasserman, (2009, February 23). How to woo 'bicultural' Hispanics. *Brandweek.com*. Retrieved from http://www.brandweek.com/bw/content_display/news-and-features/hispanic-marketing/e3icba15d3dbc5504b384e4bd5edd618afe

Wayne, T. (2009, May 4). The Hispanic audience grows online. *The New York Times*. Retrieved from <http://www.nytimes.com/2009/05/04/technology/internet/04drill.html>

Williamson, D. A. (2006, May). Hispanic youth online: Language and culture define usage. *eMarketer, Inc.* www.emarketer.com

Appendix

Interview Questions

TV

1. How many hours of TV do you watch?
2. Which programs do you watch the most on TV?
3. Are they in English, Spanish or both?
4. Is language important? If yes, why? If no, why?
5. If it is not relevant, what is relevant (important) I about the programs you watch?

Radio

6. How much radio do you listen to?
7. What are you favorite stations?
8. Are they in English, Spanish or bilingual?
9. Is language important? If yes, why? If no, why?
10. If it is not relevant, what is relevant (important) about the programs you watch?

Cell Phone

11. Do you use your cell phone to navigate the Internet?
12. Do you pay to use any of the following: IM, watch videos, download music, email, or another, (which one)?
13. Do you receive ads on your cell phone? From what companies?
14. Do you send text messages to your friends about ad campaigns or events?
15. Are you ok with receiving a few ads on your cell phone about things you like?

Internet

Mari D. González

16. How much time do you spend online? _____ per day/week/month
17. Is it in English, Spanish or bilingual?
18. What online activity consumes most of your time? [IM/chatting, downloading, shopping, watching videos, playing games, social network]
19. What do you enjoy most about the Internet? Is it in English, Spanish or bilingual?
20. What are the top 3 internet sites you visit the most? Are they in English, Spanish or bilingual?

Social Network Sites (SNSs)

21. Do you have a profile on one of these websites? Which website?
22. What is the most popular SNS among your friends? Is it in English, Spanish, or bilingual?
23. What ad do you most remember from this site?
24. Why do you like it?
25. What makes it relevant or interesting?

Interactive Advertisement

26. What is your favorite Internet ad? Why?
27. Do you get coupons, freebees, special codes, or memberships for something? How does it work?
28. Do you share it with your friends? Do you invite your friends to be part of it? How?
29. Do you download banners, wallpapers, avatars or other things from these sites?

Food and Drinks Ads

30. Do you prefer online ads in English or Spanish or bilingual?

Interactive Food & Beverage Marketing

31. What ads are popular among your Latino friends and family? Why?
32. What is your favorite online or cell phone drink ad? If there is not one you can mention a TV ad.
33. What is your favorite online or cell phone food ad? If there is not one you can mention a TV ad.
34. Is it in English, Spanish or bilingual?
35. If it is in English, what do you like about it?
36. Are there ads in Spanish that are really cool? Why?
37. Are there ads in English that are really cool? Why?
38. Are there bilingual ads that are really cool? Why?
39. If language does not make a difference (or is not relevant) what is relevant?

How to treat GUI Options in IT Technical Texts for Authoring and Machine Translation

Johann Roturier and Sabine Lehmann

Symantec Limited and acrolinx GmbH

Abstract

This paper focuses on one aspect of controlled authoring in a localization and Machine-Translation context: the treatment of GUI options, which abound in the procedural sections of IT technical documentation. GUI options are technical terms that refer to the Software User Interface. The length and complexity of GUI options is a major problem for numerous NLP tasks, including MT. GUI options which have not been identified by NLP applications typically lead to erroneous analyses of sentences. However, few authors have focused on the identification and tagging of GUI options in IT documentation. This paper delineates an approach based on a controlled language checker that benefits both the human authoring process and Machine Translation.

1 Introduction

1.1 Example of a Controlled Language Deployment

In the last three decades, publishers of technical documentation have often attempted to improve the comprehensibility of their technical source content by implementing a Controlled Language (CL). Nyberg, E., Mitamura, T., & Huijsen, W-O. (2003) give the following definition of a CL: "A CL is an explicitly defined restriction of a natural language that specifies constraints on lexicon, grammar, and style". For instance, the use of precise CL rules has proved useful to improve the usefulness of machine translation (MT) output (Rychtycky, 2002; Takako et al., 2007).

Symantec, a software publisher specialised in security and availability solutions, has been using a CL approach since 2005 to maximize the effectiveness of a third-party, rule-based MT system within its translation workflow (Roturier, 2006). To ensure that source English content complies with style guides and MT-specific rules, *Symantec* writing teams use a customised version of the *acrolinx IQ*TM checker, which is a QA tool for authoring documentation. Writers check their new XML topics using the *acrolinx IQ*TM plug-in for the *XMetal* editor to make sure that their documents comply with approved terminology, grammar, and style rules. Before translating these topics, the localisation department can examine reports to determine whether specific issues will have an impact on overall MT quality. This paper focuses on one aspect of controlled authoring: the treatment of GUI options, which abound in the procedural sections of IT technical documentation.

1.2 CL and GUI Options

The following example shows a sentence containing two GUI options ("Extract File" and "Restore from"). The example shows that capitalization is not always helpful in identifying GUI options and that the boundaries of GUI objects can be difficult to determine:

*At the **Extract File** dialog box, in the **Restore from** box, type the following where <CD-ROM drive> is the drive letter of your CD-ROM drive.*

Senellart et al. (2001) use the terms "token name" to refer to the actual button name, alert message or dialog box content, while they use the term "token identifier" for the actual GUI object with which end users interact. In this paper the terms GUI option (example: "Extract File") and GUI object (example: "dialog box") will be used to avoid any confusion. From a CL perspective, semantic tagging is sometimes mentioned as a means to remove ambiguity from source input (Bernth & Gdaniec, 2001; Nyberg et al., 2003). However, few papers have focused on the identification and subsequent tagging of GUI options in IT documentation. While Senellart et al. (ibid) focused on the identification and handling of GUI options from a translation perspective, this paper introduces an approach that benefits both the authoring process and the translation process.

1.3 Paper Structure

This paper first provides an overview of the challenges posed by GUI options when deploying NLP applications (such as a CL checker or an MT system). Section 3 discusses two of the strategies that can be used to address these issues. Section 4 describes a novel identification solution, which is concretized by the presentation of two style rules in section 5. Section 6

elaborates the results of this solution. Finally, section 7 suggests additional benefits of this approach.

2 Challenges Posed by GUI Options

GUI options not only pose challenges for NLP applications, but are also highly important from an authoring perspective. These two issues are discussed in the following subsections.

2.1 NLP Applications

During the deployment of the CL checking technology presented in section 1.1, some technical writers reported that certain rules returned false positives and created translation problems because of the presence of GUI options in source content. Examples of such problems are presented below:

Sentence length: The rule states that no more than 25 words should be used in a sentence. This rule was sometimes violated because long GUI options were handled as separate words, as shown in the following example:

*On the **Protection Manager Servers** page, select the **Enforcer** group in the **View Servers** pane and click **Edit Group Properties** to display the **Settings** dialog box.*

If the CL checker does not deal specifically with the GUI options, this sentence contains 26 tokens. However, if all of the highlighted GUI options are tokenized as single entities, the length of the sentence is reduced to 21 tokens, which prevents the false alarm for sentence length.

Grammar and style rules: False positives also occurred with regard to grammar and style rules. As an example, the grammar rule governing subject-verb agreement returned false positives because of POS tagging errors created by GUI options. As shown in the example above, GUI

options can include verbs used in unusual positions – such as "View" following an article, or "Edit" following a verb.

Term extraction: GUI options were often presented as candidate terms - in a truncated form or with the following GUI object as part of the candidate term. However GUI options should be completely excluded from term extraction.

POS assignment: Even though GUI options may contain prepositions or even punctuation marks, they correspond to a proper noun from a syntactical point of view. Ideally they should be assigned only one POS, (such as NNP from the *Penn Treebank Tagset*).

Translation: From a translation or localisation perspective, GUI options have to be translated according to the translation found in a software glossary, as the following example shows:

*To start the program, click **Start**.*

The capitalized word "Start", should be translated as "スタート" in Japanese, not "開始". In a translation workflow that uses an RBMT system, one solution might be to encode "Start" > "スタート" as a User Dictionary entry, but this entry may have side effects (for instance, if "Start" is used at the beginning of an imperative sentence). On the other hand, if the MT system is not fine-tuned to handle GUI options, the resulting translation output might be unusable, as shown in the following sentence and its corresponding French output (translated using a *Systran 5.05* engine with the Computer/Data-Processing domain enabled):

*Enter information in the **Connect to a Media Server** dialog box.*

Écrivez l'information dans le connecter dans une boîte de dialogue de serveur multimédia.

To work around these problems, contextual information is required around the GUI option to ensure that UD entries are only used in specific contexts. All these problems occur because the NLP applications do not treat GUI options as single entities. GUI options should be treated as a "unit" whatever their length is. However, this approach presupposes that the NLP applications can identify them as such.

2.2 Authoring Perspective

From an authoring perspective, GUI options should match the strings that are present in the software itself. However, software strings may be updated by software developers on a regular basis, so it is sometimes difficult for technical writers to reflect the changes in their documentation. Typical discrepancies between software and documentation include capitalization, hyphenation, and spelling. If technical writers do not extract GUI options from their documentation to conduct regular cross-references with software strings (using semantic tagging), these discrepancies are bound to go uncorrected. This ultimately affects the end-user's experience and subsequent translation process. The discrepancies may even persist when a structured XML documentation model is used (such as *DocBook* or *DITA*) because not all GUI options are necessarily tagged.

3 Treating GUI Options as Single Entities

This paper proposes a two-fold solution to treating GUI options as single entities by using semantic tagging and token class encoding.

3.1 Semantic Tagging

In the typical information development work-flow described earlier, writers must ensure that their XML topics comply with a subset of the *DocBook* XML DTD (Walsh and Muellner,

2006). The full *DocBook* DTD contains various tags related to GUI options (`guibutton`, `guiicon`, `guilabel`, `guimenu`, `guimenuitem` and `guisubmenu`). Yet to date, the only tag in use has been `<guimenuitem>`, which sets off some of the GUI options that end users have to interact with. This standard was initially designed to highlight GUI options with visual formatting in the final deliverable - usually a PDF file or a help file. The final deliverable can be obtained by transforming XML topics using an XSL-FO technology. Figure 1 illustrates this, in which the following XML sentence has been transformed:

*Click `<guimenuitem>` **Yes** `</guimenuitem>` to install the Enterprise edition.*

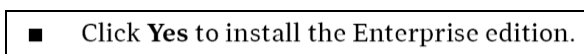


Figure 1: Rendering of a `guimenuitem` element

When this `<guimenuitem>` tag is present in the source text, an MT engine can be configured to handle it in a specific manner. For instance, the use of XSL files can achieve this objective when using certain versions of *Systran* software (Senellart & Senellart, 2005).

The next section focuses on how the *acrolinx* tokenizer handles GUI options.

3.2 GUI Options as Tokens

The *acrolinx* tokenizer has three main tasks. The first is to split incoming text into "tokens". Tokens typically correspond to words. In general, it is useful to define longer sequences as tokens if they form a meaningful unit. Examples would be compounds such as "vice versa", but also more text-specific units like GUI options. The second task is to assign "token classes" to these tokens. The token class is the name associated with the type of token. As an example, the token "Test" has the token class "FirstCapitalWord", "123" the token class "Number". Finally, the tokenizer must split the incoming text into sentences (segmentizer). In

applications like the *acrolinx IQF™* checker, which deal with huge amounts of genuine and unstructured data, the tokenizer has a prominent status. The errors occurring at that level can engender mistakes that might be visible to the technical writers. It is therefore important to treat GUI options as one token in order to ensure that the issues discussed in section 2.1 can be easily addressed.

4 Identifying GUI Options

4.1 Using Xpath and Regular Expressions

GUI options can be extracted from source documentation with a basic XPath expression (such as `//guimenuitem` or `//guilabel` if a *DockBook* DTD is used). However, untagged GUI options may require a manual (and time-consuming) extraction process if they are to be cross-referenced with software strings (as suggested in section 2.2). This manual process may, however, be alleviated if a regular expression-based approach is used, as shown by the following example:

```
((?:(?:\b[IiOo]n|[Tt]o|[C]heck|[Ff]rom|[Uu]sing)the)([A-Z][^,<>:\.]+?)\b(?:drop[-]?down list|drop[-]?down box|option|dialog box|field|drop[-]?down menu|menu|tab|box|window|dialog|check box|text box|panel?|folder|list|icon|login prompt)\b[,\.])
```

Greediness is the major problem of this approach. With this approach, the substring "Advanced tab of the Enforcer Settings" is extracted from the following string:

*On the **Advanced** tab of the **Enforcer Settings** dialog box,*

Section 4.2 discusses how a two-pronged strategy can deal with this greediness issue.

4.2 Using a Customised Solution Based on the *Acrolinx IQ™* Technology

Token class definition based on context: The tokenizer allows the definition of rules based on the context. A token class can be defined on the tokens which precede and follow the token. With the token class GUI option in mind, such an approach presupposes relatively consistent texts with respect to the use of GUI options and a good knowledge of the syntactic context in which the GUI options might occur. As shown in section 4.1, these conditions were met. The context which had initially been used for the solution based on regular expressions was also used to define the rules. A simple example of such a rule is shown below:

```
UIStringsSymantec ->  
precontext (seq (OR ("Under"), + ([WhiteSpace]))) cend,  
postcontext ([Comma]) cend,  
OR ([FirstCapitalWord], [FirstCapitalWordMixed],  
[CapitalWord], [CapitalWordWithHyphen], [OperatingSystem],  
[TrademarkSymantec]);
```

The rule expresses that the token class name is "UIStringsSymantec". The class must be preceded by "Under" and followed by a comma. The token class itself has to be one of the following token classes (which have been defined on an upper level): FirstCapitalWord, FirstCapitalWordMixed, etc. The sentence "Under Options, click OK" is then tokenized as follows:

```
Under FirstCapitalWord  
Options UIStringsSymantec , Comma  
click LowerWord  
OK CapitalWord
```

The GUI option "Options" is correctly recognized and tokenized as "UIStringsSymantec". Token class definition based on elements: The preceding approach had one significant drawback. The rules were very efficient for GUI options which were not tagged with the <guimenuitem> presented in section 2. But for the GUI options that were already tagged, the tokenization turned out to be partially redundant: why write complex rules to identify the GUI option if that information is already available via the element <guimenuitem>? However at that stage of the project, the tokenizer was not able to access any information about the element. A subsequent version of the tokenizer extended the formalism and made this information available to the rules.

With that extension, the following simple tokenizer rule was written:

```
GuiMenuItemSymantec ->  
seq((OR([LowerWord], [], ...)) : matchContext("guimenuitem"));
```

The rule expresses that the token class name is "GuiMenuItemSymantec" and that it matches any number of token classes occurring in the context <guimenuitem> (the expression in the OR parenthesis has been simplified here).

The sentence "Click <guimenuitem>Central Quarantine account</guimenuitem>" is then tokenized as follows:

```
Click FirstCapitalWord  
Central Quarantine account      GuiMenuItemSymantec
```

There is no interaction between the two token classes "UIStringsSymantec" and "GuiMenuItemSymantec": if a token is a potential match for both classes, the latter takes precedence. This ensures that only tokens with no predefined tagging are defined as

"UIStringsSymantec". In principal the token class "UIStringsSymantec" should not occur if all the GUI options were correctly tagged, but Symantec text does not always apply this element.

5 Developing two style rules

By introducing these two token classes, companies A and B solved the problems described in section 2.1: they reduced false alarms and improved the accuracy of POS assignment. Additionally, the two token classes help address two important problems with respect to the use of GUI options in IT texts. The first problem is that all of the GUI options which should be tagged with <guimenuitem> may not be actually annotated as such. The second problem is that some of the <guimenuitem> tags may be wrongly placed since the GUI objects should be outside of the element

5.1 Rule A: add_guimenuitem_tags

We decided to add a style rule to the *acrolinx IQ*TM rule formalism (Bredenkamp et al., 2000) which would mark potential GUI options where the <guimenuitem> tag is missing. However, this corresponds exactly to the content of the token class UIStringsSymantec, since this token class tries to identify potential GUI options. Developing the style rule "add_guimenuitem_tags" basically came down to marking the string of the token class "UIStringsSymantec":

```
@uistringsSymantec      ::= [ TOKCLASS "UIStringsSymantec" ];  
TRIGGER(80) == @uistringsSymantec^1  
-> ($uistring)  
-> { mark: $uistring;}
```

Rule B covered the exceptions to the rules.

5.2 Rule B: Exclude_UI_Objects

Rule B was developed based on the fact that the token class "UIStringsSymantec" sometimes slightly overgenerated due to a greedy match: UI objects such as "check box", "dialog box", "tab", "window", etc. were wrongly tokenized as part of "UIStringsSymantec". As an example:

"Inventory tab in the Properties window"

was tokenized as "UIStringsSymantec", although in fact it had to be tagged with two <guimenuitem> tags:

```
<guimenuitem>Inventory</guimenuitem> tab in the  
<guimenuitem>Properties</guimenuitem> window
```

Rule B can thus be seen as an extension of Rule A, since it flags the token class "UIStringsSymantec" (as does Rule A), but only those cases which contain UI strings.

```
@exception ::= [ TOK " (click|dialog  
box|field|icon|list|menu|OK|pane|panel|select|tab|window) "  
TOKCLASS "UIStringsSymantec" ];  
TRIGGER(80) ==  
@exception ^1  
-> ($exception)  
-> { mark: $exception;}
```

The rule makes writers aware that if they add the tag <guimenuitem>, the GUI object ("tab", "window") must be excluded.

6 Evaluation data and results

6.1 Evaluation Data

To evaluate the two rules described in section 5, we created an evaluation corpus of XML documents (in the region of 100 000 tokens). We used two sets of product documentation (pertaining to the security and availability domains) to create this corpus. These documents were not final documents. They had been developed by various product writing teams but still had to go through an editing phase. These documents contained 964 `guimenuitem` elements, following the standard usage presented in section 3.1.

6.2 Evaluation Setup

Rule A and Rule B were added to a development rule set and used with the *acrolinx IQ*TM Batch Checker to extract occurrences of untagged GUI options. Two reports were then generated and post-processed to extract contextual information for verification purposes. Due to the size of the corpus, we decided to focus on precision and use a pragmatic approach to evaluate recall: after calculating precision scores for each rule, we took a sample of the corpus to determine the types of constructs that were not found by the rule. The second part of section 6.3 describes these constructs.

6.3 Evaluation Results

Rule A flagged 311 occurrences of untagged GUI options, and Rule B reported 8 occurrences of GUI option combinations. A careful examination of the data showed that the precision of each rule was as follows: Rule A: 304/311 (97%) and Rule B: 7/8 (87%). The precision of Rule A was high (97%) despite the high number of varied structures that were found in the data – in the following examples, the GUI option appears in bold:

- Single capitalized noun: *On the **Action** menu, click **Connect**.*
- Series of capitalized nouns (including acronyms): *On the **Add DHCP Server** dialog box, enter the following information for the server that you are adding:*
- Series of capitalized nouns with unusual syntax: *lists the options in the **Media Verify** dialog box.*
- Series of lowercase words following a capitalized word: *To update the signature file, select **Download signature file**, and specify the download URL.*
- Series of lowercase words (including prepositions) following a capitalized word: *In the **Edit Server Groups** dialog box, ensure the group you want to edit is selected in the **Choose a group to edit** option.*

In this last example, the preposition "to" introduces a complement of the noun "group". Without deeper linguistic information, however, it would be difficult to identify purposive structures accurately. Rule A returned seven false positives. These can be classified into the following categories:

- Missing source words (it should read "dialog under the Advanced Settings tab"): *The number of times the user can postpone a download and for how long is controlled by the options on the **Host Integrity Setting** dialog **Advanced Settings** tab.*
- Unusual pre-modifier for the word "dialog" (the word "editor" indicates a type of dialog): *In the **Custom Requirement** editor dialog, select the node in the tree where you want to add the action.*
- Greediness: *Then, select **Incremental** from the **Backup Method** tab.*

Three other occurrences of greediness issues were found in the data. Yet this problem would not have a significant impact in a typical authoring environment because technical writers would be able to reduce the scope of incorrectly flagged GUI options. If this approach was used to transform source text in an automatic manner without human intervention, however, this problem would be more serious. The only invalid flag for Rule B occurred in the following sentence:

*To view subnodes, double-click **Jobs** or click the icon next to the node.*

This sentence was missing a standard <guimenuitem> tag around the word "Jobs", which explains why Rule B triggered in this particular context. From a recall perspective, Rule A did not catch GUI options in the following contexts:

- GUI option used as post-modifier: *The option **Enable** appears only if the robotic library is disabled.*
- GUI option's context spanning over multiple segments: *You can select all types of jobs, or any combination of backup, restore, media rotation, or utility jobs. **All**, which lets you view logs for all jobs.*
- GUI option containing two words separated with a slash: *You can remove the administration console from a Windows computer by using the **Windows Add/Remove Programs** option.*
- List of GUI options with which the user may not interact: *The **Jobs** node contains subnodes that are called **All Jobs**, **Scheduled Jobs**, **Active Jobs**, **Job Logs**, and **Schedule View**.*

- Unusual post-modifiers (such as "filter" or "node"): For example, suppose you select a certain user in the *Users filter* and **Backup** under the *Job type filter*. The *Storage Devices* node lets you view the following:

The last category shows that the rule relies heavily on the use of standard terminological modifiers. Nonetheless, a minor rule update would improve recall when dealing with texts originating from different subdomains. Overall, these results suggest that the high precision of the rule would not have a negative effect on the productivity of technical writers and could be implemented in a production environment. Section 7 will report on how writers can use these two rules in an effective manner.

7 Application and further benefits

The application of the approach presented in section 5 and evaluated in section 6 can be implemented in two ways. It can be implemented in an interactive manner within an XML editor supported by *acrolinx IQ™*, as shown in Figure 2:

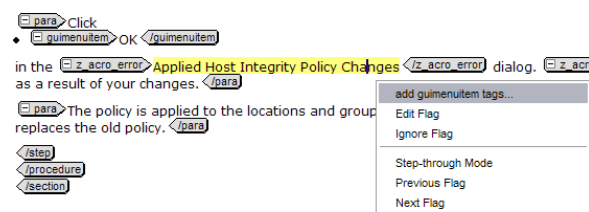


Figure 2: A violation of Rule A is reported in the *XMetal* editor

In Figure 2, a contextual menu warns the user that the GUI option "Applied Host Integrity Policy Changes" should be tagged with `<guimenuitem>`. The *acrolinx IQ™* Batch Checker can also be used to obtain a report on the number of untagged GUI options in a set of documents. This

approach is extremely useful when dealing with legacy documentation that has never been localised (to estimate how much work will be involved to handle GUI options).

While all technical writers may not necessarily see the benefits of tagging content that will not be formatted in final deliverables, the following four points should be kept in mind. First, without tagging GUI options, it is difficult to ensure consistency between software and documentation. Second, the final appearance of deliverables may change in the future. It is much easier to make an update to an XSLT file than to edit a set of source files manually. Third, translation requirements may change in the future. Again, it is easier to make an update to an XSLT file used for translation purposes than to ask translators or post-editors to spend time checking for GUI options. Regardless of whether MT is used in the subsequent translation workflow, the use of semantic tagging is useful because it also gives translators indications about the boundaries of GUI options. Finally, this approach may be seen as a first step towards the automatic tagging of content. While tags currently have to be manually inserted one by one, this will ideally change in the future to improve the productivity of writers.

8 Conclusion

This paper has presented a novel approach to identify GUI options with a view towards helping writers use semantic tagging in English technical documentation. The evaluation of this approach showed that high accuracy can be achieved when source documents are already well structured and when standard terminology is used despite the complexity of GUI options. The high accuracy of the rules should allow for a smooth introduction within a production environment, with benefits for both source consistency and translation efficiency. The rules indeed solve the challenges discussed in section 2. While this approach focused on data

Johann Roturier & Sabine Lehmann

originating from a single software publisher, it may also benefit other stakeholders from the IT industry. Software publishers rely extensively on online knowledge bases to help their customers solve problems, so having a consistent way to identify and treat GUI options improves the experience of end users.

References

- Bernth, A., & Gdaniec, C. (2001). MTranslatability. *Machine Translation*, 16(3), 175-218.
- Bredenkamp, A., Crysmann, B., & Petrea, M. (2000). Looking for Errors: A Declarative Formalism for Resource-Adaptive Language Checking. In *Proceedings of the 2nd International Conference on Language Resources and Evaluation (Greece), Athens (pp. 667-673)*.
- Habert, B., Adda, G., Adda-Decker, M., Boula de Marèuil, P., Ferrari, S., Ferret, O., et al. (1998). Towards Tokenization Evaluation. *First International Conference on Language Resources and Evaluation (LREC'98), Grenada, Spain, 427-431*.
- Nyberg, E., Mitamura, T., & Huijsen, W-O. (2003). Controlled Language for Authoring and Translation. In H. Somers (Ed.), *Computers and Translation: A Translator's Guide*. (pp. 71-110) Amsterdam: John Benjamins Publishing Company.
- Rychtycky, N. (2002). An Assessment of Machine Translation for Vehicle Assembly Process Planning at Ford Motor Company. In S. Richardson (Ed.) *Machine translation: from Research to Real Users: Proceedings of the 5th Conference of the Association for Machine Translation in the Americas, AMTA 2002 (USA), Tiburon, CA (LNAI 2499), (pp. 207-215)*. Berlin Heidelberg: Springer-Verlag.
- Roturier, J. (2006). *An Investigation into the Impact of Controlled English Rules on the Comprehensibility, Usefulness, and Acceptability of Machine-Translated Technical Documentation for French and German Users*. Unpublished PhD thesis, Dublin City University, Ireland.

Johann Roturier & Sabine Lehmann

Senellart, P., & Senellart, J. (2005, November). SYSTRAN Translation Stylesheets: Machine Translation driven by XSLT. In *Proceedings XML Conference & Exposition (USA)*, Atlanta.

Senellart, J., Plitt, M., Bailly, C., & Cardoso, F. (2001, September). Resource alignment for machine translation or implicit transfer. In *MT Summit VIII: Machine Translation in the Information Age: Proceedings (Spain)*, Santiago de Compostela (pp. 317-323).

Takako, A., Schwartz, L., King, R., Corston-Oliver, M., & Lozano, C. (2007, September). Impact of controlled language on translation quality and post-editing in a statistical machine translation environment. In *MT Summit XI (Denmark)*, Copenhagen (pp.1-7).

Walsh, N., & Muellner, L. (2006). *DocBook: The Definitive Guide*. O'Reilly & Associates, Inc.

Available online at <http://www.docbook.org/tdg/en/html/docbook.html>

The evaluation of pragmatic and functionalist aspects in localization: towards a holistic approach to Quality Assurance

Miguel A. Jiménez-Crespo

Department of Spanish and Portuguese

Rutgers, The State University of New Jersey

Abstract

Current localization QA models are based on componential error-based approaches to quality evaluation. However, as it is more complex to objectively measure pragmatic issues than language or functionality problems, the communicative-pragmatic adequacy of the target text is normally ignored while assuring that a text "look[s] like it has been developed in-country" (LISA, 2004, p. 11). Consequently, most QA models do not incorporate a category for this type of issue. This paper presents a critical review of the notion of quality in localization and its implications in current QA practices from a functionalist perspective (Nord, 1997). The main goal is to set the foundation for an evaluation process that can account for functionalist and pragmatic inadequacies through the use of localization evaluation corpora. The main issues and applications are illustrated through examples extracted from the 40,000 webpage Spanish Web Evaluation Corpus compiled by the author (Jiménez-Crespo, 2008a).

For over two decades, the production of digital content such as websites, software products, or videogames has increased exponentially. This digital revolution, in the context of a wide globalization process, has led to a dramatic increase in the amount of digital texts that are translated around the world (Cronin, 2003). The localization industry emerged as a response to this challenge, and it developed in a rapidly changing environment that demanded the translation of massive amounts of digital texts rapidly, efficiently, and cost-effectively¹. However, as some scholars have pointed out, this fast pace has meant that the industry established its practices without fully relying in the body of knowledge of related disciplines, such as Translation Studies, Interlinguistic communication, Discourse Analysis, Contrastive Rhetoric, etc. (Dunne, 2006a; Pym, 2003).²

This relative lack of dialogue between a the Localization Industry and Translation Studies has been problematic as certain established practices, such as current localization evaluation, are not based on explicitly formulated theoretical models or solid empirical research. In fact, the industry's approach to evaluation has been characterized as "experienced-based or anecdotal" (Colina, 2008), a procedural approach that, given that it is based on evaluators' previous experiences, can be subjective and, therefore, might not provide a sufficient degree of validity and reliability.

Moreover, current QA models do not implicitly incorporate the pragmatic and functional adequacy of the target text, even when this category appears in most translation error typologies

¹ The localization industry traditionally consisted of software and hardware manufacturers and the companies that produced localized versions, the so-called localization service providers. Nowadays, since the creation of the WWW, this industry also consists of telecommunication companies, web service providers and web development companies (Dunne, 2006a), as well as localization training institutions such as the Localization Research Centre at Limerick, Ireland.

² It could be argued that some other economic sectors, such as the healthcare industry in the US, have also ignored to some extent the body of knowledge of translation research.

(Melis & Hurtado, 2001). This means that QA systems might not evaluate whether the pragmatic configuration of the target text is similar to those that “have been developed in-country” (LISA, 2004). In this context, the objective of this paper is to contribute to the foundation for a QA model that can account for these recurrent inadequacies found in published localized texts. The empirical data is provided by the contrastive analysis of a 40,000 web page comparable corpus (Baker, 1995) of original Spanish corporate websites and localized sites from the largest US companies with sites addressed to customers in Spain.³ In this analysis, the notions of corpus-based evaluation (Bowker, 2001) and holistic translation assessment (Colina, 2008; Waddington, 2001) play an essential role. It will be argued that a combination of these notions could lead towards a more objective base that measurement and assessment scholars (Harper, M., O'Connor, K. & Simpson, M., 1999), as well as Translation Studies scholars (Colina, 2008; House, 1997, 2001; Melis & Hurtado, 2001;) have deemed essential for evaluation processes to be valid and reliable. It should be noted that the scope of this paper is not to propose a localization evaluation model or adapt pragmatic evaluation proposals that are generally too complex for professional settings (i.e. House, 1997), but rather to direct attention to recurring pragmatic issues in real localized texts that are difficult to detect applying current quality metrics. In doing so, a more comprehensive approach to evaluation in localization will emerge. This approach will be based on explicitly formulated theoretical models and deduced from a principled collection of texts subject to current quality evaluation practices.

³ In Translation Studies, a comparable corpus can be defined as "a structured electronic collection of texts originally written in a particular language, alongside texts translated into that same language" (Baker, 1995: 234).

Quality Evaluation in Localization

Recently, the evaluation of quality in localization has attracted the attention of a rising number of scholars from different perspectives, such as Translation Studies (Bass, 2006; Dunne, 2006b; García, 2008; Jiménez-Crespo, 2008a; Pierini, 2007;), cognitive science (Börjel, 2007) and web content management (Gibbs & Matthaiakis, 2006). These studies indicate the need for further research into this phenomenon, in part because the same set of evaluation criteria cannot be uniformly applied to all translation activity (e.g. Honing, 1998, p. 14; Larose, 1998, p. 164; Melis & Hurtado, 2001, p. 284). Nevertheless, the fact that localization is relatively young, along with the relative lack of descriptive empirical research into this phenomenon, imply that it is too early to identify a conventional set of canonized criteria for evaluating localization as a distinctive translation modality (Wright, 2006). As mentioned previously, several scholars have indicated that the industry's efforts have been carried out without fully relying on the body of knowledge of Translation Studies (Dunne, 2006a; Pym, 2003; O'Hagan & Ashworth, 2003). The lack of clear theoretical basis is somewhat problematic because, as Julianne House (1997) rightly pointed out:

Evaluating the quality of a translation presupposes a theory of translation. Thus different views of translation lead to different concepts of translational quality, and hence different ways of assessing it. (p. 7)

This implies that the notion of evaluation in localization might be somewhat different from the same notion in Translation Studies, despite that fact that an agreed upon model does not even exist in the latter.

If industry QA practices are seen in the light of translation evaluation, their goal is usually summative in nature, or in other words, they are used to determine the end result or pass a final judgment (Melis & Hurtado, 2001). The other two current theoretical approaches to

translation evaluation are the diagnostic and the formative. The main differences between these models can be found in the goals of the evaluation process: formative evaluation usually requires providing constructive feedback from the trainer to the evaluatee (Kussmaul, 1995), while diagnostic evaluation is directed towards level placing. The summative approach used by the industry is mostly directed towards passing a judgment on a pass/fail basis through an error-based assessment method. The LISA QA model is an example of this last approach, and given its dominant position in this sector, it will be reviewed later in the paper.

Any analysis of pragmatic issues has to be contextualized in the wider context of the localization industry and Translation Studies. It is therefore necessary to review current implemented practices in the industry and the underlying notions related to them, such as "quality", "errors" or the "goals" of the evaluation process.

The Notion of "Quality" in the Localization Industry

The quest towards achieving high levels of quality in all entrepreneurial processes has found a reflection in this industry, with "quality receiving more attention than ever" (Bass, 2006, p.6). Currently, it is assumed that a customizable componential QA model with an error-based acceptability threshold can guarantee a sufficient degree of quality, regardless of the genre, text type, register or display device.

If the industry's literature on this issue is analyzed, it can be deduced that most international standards define quality as the capacity to comply with a set of parameters pre-defined by the customer. For example, the ISO 9000 defines quality as: "the totality of features and characteristics of a product or service that bears on its ability to satisfy stated or implied needs" (ISO 9000, apud Ørsted, 2001, p. 443). With a very similar perspective, TQM (Total

Miguel A. Jiménez-Crespo

Quality Management) defines quality as "fully satisfying agreed customer requirements". In the same line, the definition laid out by the ASTM standard defines quality as "the degree of conformance to an agreed upon set of specifications" (ASTM, apud Melby, 2006). Nevertheless, it would be theoretically and methodologically impossible to predefine the notion of "quality" in all translated texts: for this reason, common definitions of quality usually focus on procedural aspects as opposed to establishing what could be considered a "quality" translated text.

Basically, such definitions govern procedures for achieving quality, rather than providing normative statements about what constitutes quality (Melis & Hurtado, 2001, p. 274). They are generically process-oriented instead of product-oriented (Corpas, 2006; Wright, 2006, p. 256).

As a result, the final decision about quality resides in a time-constrained evaluation process carried out by one or more evaluators that might lack the necessary theoretical framework in order to separate their own subjective judgment (House, 2001). Following previous empirical studies of translation quality evaluation (Colina, 2008; Nobs, 2006), this paper asserts that functionalist approaches could provide a valid framework to contextualize QA models.

Quality and a Functionalist View of Localization

The operative quality definitions used in the industry are consistent with functionalist approaches to translation (Holz-Mänttari, 1984; Nord, 1997; Reiss & Vermeer, 1984). In this theoretical framework, the main determining factor of the translation process is the purpose or *Skopos* of the translational action (Nord, 1997, p. 27). The source text is no longer the determinant factor in the translation or evaluation process, as the function of the target text and the context of reception are the yardsticks to which all translation decisions should be evaluated. Thus, a target text is assumed to be "functional" when it relates to the "expectations, needs,

previous knowledge and situational conditions" (ibid, p. 28) of the receiver for whom it is intended. In order to determine the functional configuration of the target text, functionalist theorists introduced the notion of "translation brief", a summary of basic aspects such as intended function(s), situational aspects, intended audience, medium, place, etc. (Nord, 1997). These aspects are assumed to help guide the translator's decisions as well as the evaluator's judgments.

If the above mentioned quality definitions are examined in more detail, they are specifically consistent with the revision of the earlier *skopos* theory by Christiane Nord (1997), who presented a "function-plus-loyalty" model. In this proposal, the translator as an intercultural communicator should not merely focus on the function of the target text in the sociocultural environment of reception, but be loyal to both the *skopos* of the translation and the requests of the initiator,⁴ as stated in the translation brief. Nevertheless, if this approach is fully implemented, the emphasis that quality definitions place on the demands of the initiator or client has to be balanced with situational and functional aspects of the context of reception. Additionally, as the most repeated criticism to functionalist approaches states, initiators might lack the knowledge and understanding of the translation phenomenon, rendering them unable to set a requested predefined notion of quality. In localization, this paper agrees with Dunne (2006b) in that:

Clients often *cannot* provide all the necessary standards, requirements or specifications for the simple reason that they are unfamiliar with the languages, culture, conventions and legal requirements of the target locale(s). (p. 100)

⁴ From a functionalist perspective, an initiator can be defined as "the person or group of people or institution that starts off the translation process and determines its course by defining the purpose for which the target text is intended" (Nord, 1997: 139).

In this sense, it is problematic to establish and measure a required level of "quality" in an objective and valid manner if different participants do not share the same notion. In localization, it has been suggested that localization researchers should direct their efforts towards educating initiators on how to achieve and measure an objective level of "quality" (Bass, 2006), but again, this could only be done once empirical research into this phenomenon can provide a clearer objective foundation. In order to establish the objective of the evaluation process and the function of the localized texts, a review of the industry literature can be revealing.

The translation "Skopos" and the Objectives of the Evaluation Process

At this point, it is logical to assume that "quality evaluation" entails a balance between a notion of quality that is soundly based on translation theory (House, 2001), empirically tested methods (Colina, 2008; Rothe-Neves, 2005), and the "objectives" set forth by the initiators that request the translation (the translation brief). The validity of the process, or in other words, "the extent to which an assessment measures what it is designed to measure" (Harper et al., 1999, p. 49), thus depends on a clear definition of what is actually being evaluated. From industry publications, it can be deduced that the objective of the localization process should be producing quality texts "with the look and feel of locally made products" (LISA, 2003, p. 5), or products "that look like [they have] been developed in-country" (LISA, 2004, p. 11). In its turn, the objective of the evaluation process would be to guarantee that these texts do look like local productions and serve their purpose efficiently. Therefore, it is fair to argue that, apart from eradicating any transfer and language errors in the target language, a functional target text should "look" like texts originally produced in the target region. In order to accomplish this goal, compliance with existing conventions in natively produced texts in the target culture becomes an

essential aspect of quality (Jiménez-Crespo, 2009; Nielsen & Tahir, 2002).⁵ Conventions are shared by a discourse community and frame the expectations from receivers and possible effects of the texts on them. Conventions are culture-dependent and may differ from culture to culture (Nord, 1997, p. 54); this is the reason why they play an important role during translation processes. In localization, target texts are presented as original productions, that is, receivers do not need to be implicitly aware that they are interacting with a translated text. Their main goal is to produce in the reader an effect as close to that of a reader of the original text (Newmark, 1981, p. 39), a type of translation that has been referred to as instrumental (Nord, 1997), covert (House, 1997) or communicative translation (Newark, 1981). Thus, translators in principle would replace any conventional feature in the original text with the conventional feature in the same textual genre in the target text. For the purposes of QA analysis, this means that interlinguistic studies centered on describing conventional features in different digital genres and locales could be highly beneficial (Jiménez-Crespo, 2009, 2008a; Adkisson, 2002; Nielsen & Tahir, 2002). As an alternative approach, companies such as *Facebook*, *Microsoft* or *Symantec* have already implemented novel ways to identify conventional features shared by target discourse communities. This is the case of "crowdsourcing", or in other words, letting communities of web users decide and vote on the best translations for terms and phrases (O'Hagan, 2009).

Once the notion of quality has been reviewed from a functionalist perspective, the next step would be to review the ill-defined notion of "error". Functionalist researchers normally downplay the importance of errors in any evaluation process, as they assert that error analysis per-se is insufficient to evaluate the quality of a translation. Instead, "it is the text as a whole

⁵ The notion of convention can be defined as: "Implicit or tacit non-binding regulation of behavior, based on common knowledge and the expectations of what others expect you to expect from them (etc.) to do in a certain situation" (Nord, 1991: 96).

Miguel A. Jiménez-Crespo

whose function(s) and effect(s) must be regarded as the crucial criteria [...]" (Nord, 1991, p. 166). Nevertheless, it is obvious that errors play an essential role in QA models and will continue to do so. A review of this concept is therefore needed.

The Role of Errors in Quality Evaluation

Errors and error typologies are the foundation upon which the QA evaluation processes are built. Normally, they include pre-established error typologies, but again, a clear and operative notion of error is not provided. In Translation Studies, the notion of error has been defined from a cognitive and functionalist perspective as a translation problem that has not been solved or that has not been solved appropriately (Nord, 1996, p. 96-100). Translation errors are therefore related to translation problems⁶, and they can be defined as "an objective problem which every translator [...] has to solve during a particular translation task" (Nord, 1991, p. 151). After a thorough review of the notion of error in previous literature, Hurtado Albir (2001, p. 305-306) categorized them according to three possible principles: (1) errors caused by a misunderstanding of the source text, such as false sense, omission, no sense, etc., (2) language errors in the target language, such as punctuation, grammar, lexical or style errors, (3) and pragmatic errors, such as those related to inadequacies as far as the function or *skopos* of the translation is concerned (Reiss & Vermeer, 1984). These pragmatic errors can only be identified by "comparing the source and target text in the light of the translation brief" (Nord, 1997, p. 76), and native speakers reviewing exclusively the target text might not be able to identify them. This is due to the fact that pragmatically erroneous expressions might not be inadequate in themselves; they only become inadequate "with regard to the communicative function it was

⁶ Translation problems are also associated to translation strategies. From a cognitive perspective a translation strategy is defined as "a potentially conscious procedure for the solution of a problem which an individual is faced with when translating a text segment from one language to another" (Lörcher, 1991: 76).

supposed to achieve" (ibid, p. 73). These are normally more serious errors, as receivers do not realize that they are getting wrong or inadequate information. This contrasts with general language or cultural errors that can be more easily identified and corrected (Colina, 2008). Current localization error typologies do not explicitly incorporate a category for pragmatic inadequacies, and as such, QA and QC processes cannot trace the actual cause of the error and try to put a mechanism in place to correct it (Bass, 2006). It is our contention that localization QA models are in dire need of adding this third category. This could be implemented through a combination of corpus-based approaches to evaluation, the training of evaluators and/or the adoption of a holistic system (Colina, 2008; Waddington, 2001).

In order to illustrate these types of errors, the following list presents recurrent pragmatic inadequacies in published localized web texts into Spanish. It was extracted from the multilevel corpus analysis carried out by the author in his doctoral dissertation (Jiménez-Crespo, 2008a, p. 536-566). These errors were found in a corpus of 95 published localized websites from the largest US companies according to the Forbes list. These sites were contrasted with 172 original Spanish corporate sites following a corpus-based methodology:

1. More direct formulation of the appellative function in Spanish localized texts. As an example, the concordance analysis of the verb "leer" 'read' shown in Figure 1 illustrates how the infinitive form of the verb is more frequent in Spanish original privacy pages to address the user, while imperative form of this verb, such as "lea/lee" is overwhelmingly used in localized texts (81.81%).

| Original Spanish Corpus | | Localized Spanish Corpus | |
|-------------------------|---|--------------------------|--|
| N | Concordance | N | Concordance |
| 1 | Usuario también ha de leer atentamente las co | 1 | e desde este sitio web. Lea nuestra Declaració |
| 2 | uencia, el Usuario debe leer atentamente el pres | 2 | ia Rogamos lea atentamente esta inf |
| 3 | uario se compromete a leer atentamente las con | 3 | Personales. Por favor, lea atentamente esta pol |
| 4 | Usuario también ha de leer atentamente las cor | 4 | tes de utilizar este sitio lea atentamente estos t |
| 5 | rvicios, el usuario ha de leer atentamente tambié | 5 | de tales datos. Por favor lea cuidadosamente est |
| 6 | LE ROGAMOS QUE LEA CON ATENCIÓN L | 6 | l Avisos legales Lea cuidadosamente los |
| 7 | uario se compromete a leer detenidamente el co | 7 | ormación en este sitio, lea detenidamente esta |
| 8 | recomienda al usuario leer el mismo atentamen | 8 | ===== *ATENCIÓN: LEA DETENIDAMENTE |
| 9 | e Aviso Legal, debiendo leer el Usuario atentame | 9 | rvicios. Le rogamos que lea detenidamente las p |
| 10 | alga de este Sitio Web, lea la política de privacidad | 10 | estas comunicaciones, lea el punto 10 a contin |
| 11 | ue los usuarios deberán leer las mismas cada v | 11 | o aquello que Ud. vea o lea en el Sitio Web tien |
| 12 | es utilizadas no pueden leer los archivos cookie | 12 | e uso *POR FAVOR, LEA ESTAS CONDICIO |
| 13 | es utilizadas no pueden leer los archivos cookie | 13 | l "Servicio"). Por favor, lea este Acuerdo con at |
| 14 | izada si el interesado no lee y acepta las con | 14 | *¡MUY IMPORTANTE! LEA ESTE ACUERDO L |

Figure 1. Extract from a KWIC (Key Words in Context) display of the results of a query of the verb "leer" in original and localized Spanish web privacy policies.

This illustrates that pragmatically, Spanish texts tend to use verbal periphrasis to formulate this recurring unit of meaning. Consequently, the frequent use of direct imperative forms of verbs in texts localized into Spanish could be considered a pragmatic inadequacy, such as "¡MUY IMPORTANTE! LEA ESTE ACUERDO LEGAL [sic] 'Very important! Read this legal agreement!'. As the concordance lines shown in Figure 1 demonstrate, the user is rarely addressed directly in the imperative form of the verb in original Spanish texts, but rather with verbal periphrasis "le rogamos que lea el siguiente acuerdo" 'we encourage you to read the following agreement' (Jiménez-Crespo, 2008a). In these cases, even when the translated utterance would not be erroneous per-se, the target text would not comply from a pragmatic perspective with the expectations of members of the target discourse community.

2. Not complying with sociocultural norms: 33% of localized contact forms in the corpus under study use the singular form of the word 'last name', "apellido", even when Spanish users always identify themselves with two last names, "apellidos" (Jiménez-Crespo, 2008a).

3. Explicitation in brief textual segments that need to be concise due to space constraints (Jeney, 2007). The translation of "contact us" for longer phraseological units such as "Póngase en contacto con nosotros" - "Entre en contacto con nosotros" 'get in touch with us', although correct in Spanish, would be inadequate as they do not comply with the specific constraints for this text type.

4. Using non-conventional linguistic forms associated with the translated textual genre. A specific textual genre and each of its thematical units might incorporate recurring conventional linguistic forms (Glässer, 1990; Göpferich, 1995; Gamero, 2001), such as "contact us" or "about us" in web navigation menus (Nielsen & Tahir, 2002). A descriptive study of conventional terminology in navigation menus was carried out in both components of the evaluation corpus (Jiménez-Crespo, 2009). It provided a percentage of use for all lexical units found in original navigation menus. As an example, the term "contacto" 'contact' was found to be the most conventional term in original navigation menus (49.9%), while in English texts "contact us" appears in 89% of websites (Nielsen & Tahir, 2002). In original websites, the average frequency of the most used lexical unit represented in navigation menus was 58.26%. The average frequency for the most used lexical units in original Spanish texts in navigation menus, such as "contacto" or "aviso legal", is 34.36% in localized texts (Jiménez-Crespo, 2009).⁷

After an initial research stage, these types of pragmatic errors could be incorporated to QA systems. But how can they be incorporated in current established systems such as LISA's? A review of the role of errors in the LISA QA model can shed some light into this issue.

⁷ See Jiménez-Crespo (2009) for a complete analysis.

Errors and Pragmatic Issues in the LISA QA Model

The LISA QA analysis is the most widely used evaluation model in the localization industry (Lommel, 2008). In this error-based system, the evaluator(s) set(s) a threshold for the number of combined minor, major and critical errors that are allowed, and if that pre-defined threshold is passed, the localized text would be deemed "unpublishable". Nevertheless, what counts as a minor, major or critical error is rarely described, an aspect that can significantly increase the variability among raters. This approach also presumes that a limited number of errors is unavoidable due to the economic context in which this process is performed (Bass, 2006).⁸ In addition, even when this type of evaluation can be defined as summative in nature, it also incorporates a formative component given that evaluators might be requested to provide feedback to translators in order to minimize future errors (Esselink, 2000), or they might be requested to propose solutions to all identified inadequacies (Chandler, 2004).

As with any localization evaluation, the LISA QA model incorporates both a linguistic and a functionality testing stage (Esselink, 2000; Chandler, 2004).⁹ Both phases are usually performed by different evaluators or testers.¹⁰ Nevertheless, these stages cannot be understood in isolation, and even when the linguistic testing might measure an outstanding level of quality, the overall localized text would not achieve this degree of conformance to a quality standard if it contains functionality problems, such as deficient text integration or textual segments in other

⁸ Nevertheless, it is assumed that a process should be in place to minimize and effectively correct those errors.

⁹ Linguistic testing deals with all the textual aspects of the evaluation process, and it could be compared to traditional translation testing. It can be defined as a stage that "focuses on all language-related aspects of a localized application" (Esselink, 2000: 150). Functionality testing deals with all the programming and functional aspects of a software program or website, such as the presence of programming errors, the inclusion of other languages in a localized version or that specific character encodings are correctly displayed.

¹⁰ In software testing Esselink (2000: 145-155) identifies internationalization testing, localization testing, functionality testing and delivery testing stages. Traditional translation evaluation would take place during the localization testing stage. Another study identifies linguistic, functional, visual and technical QA stages (Corte, 2000).

languages (Bass, 2006, p. 92). This is an essential aspect that distances localization evaluation from other translation evaluation processes and, given that translation evaluation has to incorporate contextual aspects (Sager, 1989), measurements have to account for both interrelated stages.

In the context of a time and resource limited process (Wright, 2006), the tension between technological and linguistic aspects of this type of evaluation has led the industry to place a higher emphasis on functionality testing, disregarding more complex linguistic aspects beyond the sentence or microtextual level of analysis (Dunne, 2006a)¹¹. In fact, out of eight componential evaluation areas in the LISA QA model, only one deals with "language", in which an error can be identified as a mistranslation, an accuracy error, terminology error, language error, style error, country or dialect error, and consistency error. Given that these errors do not have an empirical foundation; it is easy to perceive the difficulty in assigning any given error to only one of these categories. Some of them overlap, such as accuracy and style errors, and language and style errors. Moreover, consistency is the only category that could be assigned to some of the more serious inadequacies at macrotextual or superstructural levels, such as using different terms for the same section of a website. Thus, while these inadequacies are theoretically accepted to be more important (Larose, 1998; Williams, 2003), translators and evaluators are somewhat forced into a microtextual level that has difficulties in dealing with problems at the suprasegmental level, or in other words, those that affect the text as a global unitary entity. The same can be said of the more important pragmatic errors (Nord, 1997), as the

¹¹ The researcher points out: "Because functional aspects of a program can be objectively assessed [...], people often perceive software development and localization as processes that are akin to manufacturing, while forgetting about the linguistic aspects of the program and the often subjective nature of linguistic quality definitions" (Dunne, 2006b: 96).

Miguel A. Jiménez-Crespo

decontextualization of textual segments hinders the evaluation of a global coherent text from a functionalist and pragmatic perspective.

The LISA QA Model and Functionalist Approaches to Evaluation

Taking into consideration that this model responds to economic and time constraints, one of the positive aspects of the LISA QA model resides in the fact that it is componential, or in other words, that it recognizes many aspects of quality (Colina, 2008). This approach could be easily tailored to the notion of context dependent quality (Sager, 1989). It also implicitly accepts that the search for an absolute standard of quality is impossible. In this case and from a functionalist perspective, the localized text will be more or less appropriate for the purposes for which it is intended (Nord, 1997). However, the fact that the evaluator has to constantly decide on assigning weight to errors does not provide the sufficient degree of reliability and validity as stated by a number of researchers (Williams, 2003; Melis & Hurtado, 2001; Colina, 2008). Additionally, this system follows a structural and componential approach based on the identification of a customizable number of language and transfer errors, even when empirical error typologies of errors have not been fully developed. Research in multiple language combinations could produce more reliable and efficient typologies that could incorporate recurrent pragmatic, functional and communicative inadequacies. These types of pragmatic and functional features are more difficult to evaluate than grammar, meaning or spelling (Colina, 2008, p. 125) and, therefore, they tend to be absent in error-based systems such as LISA's. These pragmatic/functionalist areas have been recently incorporated through holistic evaluation proposals (Waddington, 2001; Colina, 2008) that, among others, provides for a way of evaluating the communicative/pragmatic adequacy of a localized text, "rather than being limited

to grammatical and or stylistic errors at the sentence level, and changes in meaning" (Colina 2008, p. 107).¹²

The holistic review evaluates the pragmatic/functionalist adequacy of the target text as a whole (Nord, 1997; Waddington, 2001), providing a more reliable assessment than what is possible through evaluating single segments alone. It uses descriptive statements that simplify the evaluation process, allowing raters to concentrate on specific inadequacies in the middle of a quality continuum, rather than correct/incorrect judgments. The statements used are attached to a numeric scale and they are implicitly formulated, such as the following example from Colina (2008):

1. Disregard for the goals, purpose, function and audience of the text. The text was translated without considering textual units, textual purpose, genre, audience, etc.

[...]

4. The translated texts accurately accomplishes the goals, purpose, (function: informative, expressive, persuasive) set for the translation and intended audience (including level of formality). (p. 129)

As mentioned previously, the development of an empirical error typology that would incorporate recurring pragmatic inadequacies, combined with an additional holistic level that could assess the entire text through descriptive statements, would evaluate more reliably whether texts look like they have been developed "in-country". In the development of these typologies and the holistic evaluation of the pragmatic adequacy, corpus based approaches to evaluation can provide a solid empirical foundation. The following section reviews these approaches, as well as proposes different practical ways to incorporate corpora into the evaluation of pragmatic and functionalist issues.

¹² These methods used descriptive statements in order to describe the level of adequacy of any given translation to the intended function, genre, and text type of the target text. These holistic models have also been empirically proved to be as valid and reliable in summative evaluation processes as error based approaches (Waddington, 2001).

Corpus-based approaches to quality evaluation

Corpus-based approaches to translation research have been extremely productive during the last decade (Baker, 1995; Laviosa, 2002; Olohan, 2004). The success of this approach is due to the fact that, if the textual selection process is done according to established principles (Sinclair, 1991; Biber, 1993) it provides quantitative data that can support valid and objective decisions while translating or evaluating (Shreve, 2006, p. 311). This more objective and empirical approach to evaluation has been defended by numerous scholars (Nord, 2003; House, 2001; Melis & Hurtado, 2001; Bowker, 2001), and it is somewhat striking that little research on corpus-based evaluation in localization has been done so far. Among others, Bowker (2001, p. 346) defends the importance of using corpora in evaluation, since it "entails making judgments about appropriate language use, [and] it should not rely on intuition, anecdotal evidence or small samples". It could be argued that the same rationale would apply to localization evaluation, given that corpus approaches can provide the empirical data that can help overcome the subjective bias unavoidably present in all evaluation processes (Hönig, 1998, p. 14). Using corpora could make this process much more transparent, and it would allow evaluators to make decisions that can be easily backed up by representative quantitative data (Baker, 2004, p. 184). In this case and among different corpus types (Laviosa, 2002, p. 34-38), evaluation corpora can constitute a source of helpful conceptual and linguistic information (Bowker, 2001).

Few translation scholars have applied corpus-approaches to localization. Its practical use during professional translation and evaluation has been both supported (Shreve, 2006; Jiménez-Crespo, 2009, 2008a) and rejected (Biau & Pym, 2006). The main ground for rejecting the use of corpora is the fact that they are not cost-effective when applied to professional translation

(ibid, p. 18). While time and cost issues are central in professional summative evaluation, it is fair to defend that a bottom-up approach, compiling extensive comparable or translational corpora made up of original texts in the target language, and translated texts into that same language from English,¹³ could prove to be part of a more objective base for localization evaluation. The potential uses of these evaluation corpora are described in the next section.

Uses of Corpora in Translation Evaluation

The use of corpora in Translation research, training and professional practice has been widely explored during the last two decades (e. g. Laviosa, 2002; Olohan, 2004). Many of the established principles and procedures of this methodology can be highly beneficial to localization evaluation both in the preliminary stages and during the evaluation process. The following list briefly describes the main applications of this methodology:

1. Descriptive monolingual and bilingual genre-based grammars or contrastive studies.

As mentioned above, the industry has as an objective to release products that look like non-translated texts in any give locale (LISA, 2004, p. 11), and therefore, quality should be mostly related to the expectations of the receivers (Jiménez-Crespo, 2009; Nobs, 2006). Thus, the logical point of departure for corpus-based approaches to localization evaluation entails descriptive studies of conventions established in non-translated textual genres. This could be achieved by compiling evaluation corpora of the most recurring original and localized digital genres. In the case of bilingual corpora, contrastive analyses of original and localized sites can identify linguistic features that distance localized texts from the textual prototype shared by the target discourse community (i.e. superstructural differences, stylistic and terminological

¹³ English is the most frequent source language in localization (Cronin, 2003). Nevertheless, the Internet has multiplied the localization processes in the opposite direction, from any language into English, and between multiple languages in contact in specific areas, such as Italian, German, French, Rumantz and English in Switzerland.

Miguel A. Jiménez-Crespo

variation, pragmatic differences, typology of recurrent translation and localization errors, etc.), and therefore, that could influence the receiver's quality judgment on the text. These descriptive studies can also shed some light on most functional and pragmatic inadequacies in localized texts.

2. *Evaluation corpora* (Bowker, 2001). The need for comparative empirical data in corpus studies is well known and documented (Stubbs, 1996; Kenny, 2001). Among different corpus types, a specialized control/evaluation corpus can be used as a benchmark against which the uses of any given term, collocation, phraseological unit, etc. can be quickly contrasted during a translation evaluation process. This is mostly needed in cases where the evaluator requires quantitative data in order to support or reject a possible decision. The process entails the use of a lexical analysis tool (i.e. *Wordsmith* tools, *Paraconc* or *Monoconc*) that provides applications such as concordancers, wordlists or keyword analysis. A concordancer can retrieve all instances of use of any given word or search pattern in its immediate context, and displays these in a clear format that can be reorganized according to different criteria.¹⁴ These concordancers allow wildcard searches in order to identify and analyze in context different uses of a single word, such as "virus*" in order to retrieve all word forms for this lemma such as "virus" and "viruses". As an example, Figure 2 presents a concordance line for the noun "virus" both in the original and the localized corpus of the Spanish Web Evaluation Corpus. This search was carried out in order to observe the adequacy of a translated phraseological unit in a web privacy policy, "libre de virus" 'free of viruses':

¹⁴ A concordancer can be defined as a "tool that retrieves all the occurrences of a particular search pattern in its immediate contexts and displays these in an easy-to-read format" (Bowker, 2002: 53).

| LOCALIZED CORPUS | | ORIGINAL CORPUS | |
|------------------|---|-----------------|--|
| N | Concordance | N | Concordance |
| 17 | vío constituye ni contiene virus informáticos, propag | 9 | s por la introducción de virus . Caja Madrid no se |
| 18 | mpletos y están libres de virus u otros componente | 10 | ios. - La existencia de virus , programas malicios |
| 19 | (ES) ESTÉN LIBRES DE VIRUS U OTROS COMP | 11 | ios. • La existencia de virus , programas malicios |
| 20 | este Sitio estén libres de virus ni de ningún otro có | 12 | l Usuario la ausencia de virus , gusanos, caballos |
| 21 | OR ESTÉN LIBRES DE VIRUS INFORMÁTICOS.* | 13 | ventual existencia de virus u otros elementos n |
| 22 | té absolutamente libre de virus . Antes de bajar las | 14 | garantiza la ausencia de virus u otros elementos e |
| 23 | a estará siempre libre de virus , y no aceptará ning | 15 | NTO; (v) la presencia de virus u otros elementos |
| 24 | eb de GLASER libre de virus , GLASER no garanti | 16 | garantiza la ausencia de virus u otros elemento |
| 25 | oducir o difundir en la red virus informáticos o cuale | 17 | garantiza la ausencia de virus u otros elementos |
| 26 | ducir o difundir en la red virus informáticos u otros | 18 | r previo, la ausencia de virus o elementos en los |
| 27 | pueda introducirse algún virus en el Sitio Web me | 19 | r previo, la ausencia de virus o elementos en los |
| 28 | os por la introducción de virus y/o ataques infor | 20 | controlar la ausencia de virus o cualquier otro ele |
| 29 | ir inadvertidamente algún virus informático a través | 21 | berse a la presencia de virus o a la virus o a la pr |

Figure 2. Extract from a KWIC (Key Words in Context) display of the results of a query of the noun "virus" in original and localized texts into Spanish.

The data provided by the concordance analysis shows that the phraseological unit "libre de virus" does not appear in original texts in order to formulate this recurring unit of meaning. It can be thus recognized as a calque of the original English collocation "free of viruses", and therefore, evaluators can quickly establish that this is a pragmatically inadequate expression in the target locale. As a solution, in original Spanish texts this noun is used mostly in combination with the words "presencia" 'presence' or "ausencia" 'absence'.

3. *Word frequency lists.* Frequency lists can be used in order to identify possible patterns of error during the translation evaluation process. These lists allow evaluators to discover how many words there are in a corpus (tokens) and how often each of them appears (types), thus providing evidence about recurring lexical or syntactic patterns of interference. Lexical analysis tools such as Wordsmith Tools can provide lists not only for single words, but they can also provide lists of word clusters that co-occur in a given corpus. These lists can also help spot recurring terms or phraseological units that, due to their high frequency, can be identified as borrowings or calques, such as "términos y condiciones" 'terms and conditions' or "Acerca de

Miguel A. Jiménez-Crespo

nosotros" 'About us' in localized websites into Spanish (Jiménez-Crespo, 2008a). Frequency lists for single words or clusters of words in a given text can also be compared with a tool such as Keywords with the frequency list obtained in the original or control section of the evaluation corpus. These tools compare the frequency of a single word or cluster of words in terms of percentage of use in the overall corpus, and once the frequency list of a text to evaluate is contrasted with the list of the original evaluation or control corpus, the list obtained can show recurrent lexical and syntactic patterns that are overrepresented in contrast to original texts. This methodology can help spot cases of lexical or syntactic interference in localized texts. As an example, Table 1 consists of a keyword analysis of localized legal texts in websites using the Spanish original websites as a control corpus. Once proper names were eliminated, the analysis shows that the frequency of the pronoun "usted" 'you' is six times higher in localized texts. This is due to the fact that pronouns that function as subjects in Spanish can be omitted, while they are compulsory in English. Additionally, the word "leyes" [laws] is also more frequent in localized texts. This higher frequency can be traced back to the fact that in the Spanish legal texts the concept "laws" is expressed by the term "legislación", thus the higher frequency of the word "leyes".

Table 1.

Keyword Analysis of Localized Legal Web Texts in the Evaluation Corpus Using the Subcorpus of Original Legal Web Texts as a Control Corpus. Names of Corporate Businesses were Eliminated.

| N | word | freq.loctxt | % | freq.natleg~1.txt | % | keyness |
|-----------|--------------------|-------------|-------------|-------------------|-------------|--------------|
| 2 | usted | 433 | 0,39 | 36 | 0,06 | 183,3 |
| 3 | sitio | 1.212 | 1,08 | 283 | 0,49 | 167,3 |
| 8 | información | 1.195 | 1,06 | 327 | 0,57 | 114,8 |
| 13 | privacidad | 317 | 0,28 | 61 | 0,11 | 60,7 |
| 17 | términos | 251 | 0,22 | 46 | 0,08 | 51,6 |
| 22 | línea | 55 | 0,05 | 0 | | 45,6 |
| 32 | declaración | 85 | 0,08 | 5 | | 42,7 |
| 41 | marcas | 202 | 0,18 | 41 | 0,07 | 35,6 |
| 42 | utilizar | 214 | 0,19 | 45 | 0,08 | 35,6 |
| 43 | leyes | 103 | 0,09 | 12 | 0,02 | 34,4 |

This same type of analysis can be performed with word clusters of different length, normally up to eight words, and this could provide evidence of recurring syntactic, stylistic or lexical constructions that might be inappropriate when contrasted to the control corpus of original texts.

4. *Identifying patterns of pragmatic error.* Quantitative analysis of the translational component of the corpus could be useful in establishing and identifying patterns of errors, as well as classifying them (Jiménez-Crespo, 2008b; López & Tercedor, 2008; Melis & Hurtado, 2001). As previously mentioned a classification could help establish a more reliable objective foundation during error judgments in general or while assigning a certain weight to any identified error. Additionally, a classification of the most recurring errors can direct the attention of evaluators towards more serious error patterns, such as pragmatic and superstructural inadequacies that are more difficult to identify and, therefore, require higher levels of competence or specific training from the evaluators (Williams, 2003). As an example, the analysis previously presented with the verb "leer" 'to read' in the translational component of the

Evaluation Corpus can be revealing. In Figure 3, the concordance lines associated with the verb "leer" in the localized corpus show differences in its grammatical use and in its patterns of lexical co-occurrence if compared to the original Spanish texts in Figure 1.

| N | Concordance |
|----|---|
| 1 | . Envío de información personal Lea la <u>Política de intimidad</u> de C |
| 2 | o Legal *¡MUY IMPORTANTE! <u>LEA ESTE ACUERDO LEGAL A</u> |
| 5 | s con la mayor seguridad posible. Lea la <u>política de confidencialidad</u> |
| 9 | d de AOL. <u>Le recomendamos que</u> lea la <u>normativa de privacidad</u> de |
| 14 | vigentes. <u>Le agradecemos que</u> lea nuestra Política de Privacidad |
| 15 | General Avisos legales Lea <u>cuidadosamente</u> los términos |
| 17 | desde Mayo de 2005 Por favor, <u>lea y revise atentamente</u> el acuer |
| 32 | ===== *ATENCIÓN: <u>LEA DETENIDAMENTE ESTAS C</u> |
| 35 | l a Electronic Arts, <u>asegúrate de leer</u> la normativa de privacidad de |
| 40 | ios, el Usuario también <u>habrá de leer</u> atentamente tanto este Aviso |
| 48 | s sitios. <u>Le advertimos que</u> debe leer las declaraciones de |
| 49 | B (EL "SITIO WEB") <u>ROGAMOS LEER DETENIDAMENTE LAS SI</u> |

Figure 3. An extract from a KWIC display of the results of a query of the verb "leer" 'read' in localized texts. All inadequacies or pragmatic errors are underlined.

All items underlined can be considered language errors, pragmatic inadequacies or non-conventional items in original Spanish texts. For example, the first line shows a translation of the lexical unit "privacy policy" that has been rendered as "intimacy policy". These concordance lines also show the terminological variation in the translation of "privacy policy", "política de intimidad", "política de privacidad", "normativa de privacidad", "política de confidencialidad". Furthermore, it can be observed that, while the direct appeal to the user is rarely present in original Spanish legal web texts and he/she is addressed impersonally (see Figure 1), in most localized concordance lines the user is addressed directly: "lea cuidadosamente..." 'read carefully...', "le recomendamos que lea..." 'we recommend you to read...', "le agradecemos que lea..." 'we would appreciate it if you would read...', "asegúrate de leer..." 'make sure you read...'. '.

Moreover, a contrastive genre-based study can link all identified errors or inadequacies to possible solution(s) based on the range of natural possibilities that appear in original texts. For example, the translation of the lexical unit "case studies" or "success stories" used in navigation menus in corporate websites could be linked to those lexical units that serve a similar function in original texts, such as "experiencia de la empresa" 'Company experience' o "nuestros clientes" 'our clients' (Jiménez-Crespo, 2008a, p. 369).

Conclusions

In interactive texts such as web pages or software products, the pragmatic adequacy of the localized texts is directly linked to the success of the communicative interaction established between the digital text and the receiver. Once established that the goal of the localization industry for localized texts is to look like "locally-made products" (LISA, 2003, p. 5), this paper has defended that QA systems should go beyond structural approaches centered on the identification of lexical, syntactic and transfer errors and incorporate an additional pragmatic level. Pragmatic inadequacies are harder to identify by native speakers than lexical or grammar errors: they are not erroneous per se in the target language system, but more or less inadequate with regard to the communicative purposes intended (Nord, 1997). In order to add this additional level, this paper has highlighted the importance of providing a theoretical and methodological base in order to successfully assess this integral aspect of translation evaluation.

As previously mentioned, the scope of this paper was not to propose a novel evaluation model or apply existing complex theoretical proposals, but rather, to set in place this much needed theoretical and methodological foundation as a first step towards a more valid and

Miguel A. Jiménez-Crespo

reliable QA model. Given that localization is an industrial process with limited time and resources (Sager, 1989, p. 91), the contribution of this paper can be summarized in a proposed triple approach that can complement current existing models: (1) the identification of pragmatic errors through carefully constructed representative corpus studies in order to include them in QA error-based systems (Jiménez-Crespo, 2008a), (2) the use of evaluation corpora during the assessment processes (Bowker, 2001), (3) the possible addition of a global holistic system as an additional step that incorporates a module for the pragmatic and functionalist adequacy of the target text (Colina, 2008; Waddington, 2001).

Finally, this paper has highlighted the need for more theoretical and applied research into localization in general and into evaluation in particular, as the fuzzy definitions of operative notions, such as "quality", "error", "language error", "error typology", "reliability" or "evaluation objectives", hinder to some extent the development of solid QA systems that can produce more objective valid and reliable outcomes. It is hoped that this paper will provide a starting point for establishing an agreed-upon evaluation model that can benefit not only practitioners and researchers, but also the end-users of localized digital texts.

References

- Adkisson, H. (2002). *Identifying De-Facto Standards for E-Commerce Web Sites*. Unpublished MS Thesis, University of Washington, WA.
- Baker, M. (2004). A corpus-based view of similarity and difference in translation. *International Journal of Corpus Linguistics*, 9 (2), 167-193.
- Baker, M. (1999). The Role of Corpora in Investigating the Linguistic Behaviour of Professional Translators. *International Journal of Corpus Linguistics* 4 (2), 281-298.
- Baker, M. (1995). Corpora in Translation Studies: An overview and some suggestions for future research. *Target* 7, 223-243.
- Baker, M. (1993). Corpus Linguistics and Translation Studies: Implications and Applications. In M. Baker, G. Francis and E. Tognini-Bonelli (Eds.), *Text and Technology: In Honour of John Sinclair* (pp. 233-250). Amsterdam-Philadelphia: John Benjamins.
- Bass, S. (2006). Quality in the Real World. In K. Dunne (Ed.), *Perspectives on Localization* (pp. 69-84). Amsterdam-Philadelphia: John Benjamins.
- Biber, D. (1993). Representativeness in corpus design. *Literary and Linguistic Computing*, 8 (4), 243-257.
- Börjel, J. (2007). *Language-specific quality issues in a real world localization process*. Unpublished MS Thesis. Linköpings Universitet, Sweden.
- Bowker, L. (2002). *Computer Aided Translation Technology: A Practical Introduction* Ottawa: University of Ottawa Press.
- Bowker, L. (2001). A Corpus-Based Approach to Translation Evaluation. *Meta* 46 (2), 345-363.
- Chandler, H. M. (2004). *The game localization handbook*. Hingham, MA: Charles River Media.

Miguel A. Jiménez-Crespo

- Corpas Pastor, G. (2006). Translation Quality Standards in Europe: An Overview. In E. Miyares Bermúdez & L. Ruiz Miyares (Eds.), *Linguistics in the Twenty First Century* (pp. 47-57). Cambridge: Cambridge Scholars Press/Santiago de Cuba: Centro de Lingüística Aplicada.
- Colina, S. (2008). Translation Quality Evaluation: empirical evidence from a functionalist approach. *The Translator* 14 (1), 97-134.
- Cronin, M. (2003). *Translation and Globalization*. London: Routledge.
- DIN 2345 Edition: 1998. *Übersetzungsaufträge*.
- Dunne, K. (2006a). A Copernican Revolution. In K. Dunne (Ed.), *Perspectives on Localization* (pp. 1-11). Amsterdam: John Benjamins.
- Dunne, K. (2006b). Putting the cart behind the horse: rethinking localization quality management. In K. Dunne (Ed.), *Perspectives on Localization* (pp. 95-117). Amsterdam: John Benjamins.
- Esselink, B. (2000). *A Practical Guide to Localization*. Amsterdam-Philadelphia: John Benjamins.
- Gamero Pérez, S. (2001). *La traducción de textos técnicos*. Barcelona: Ariel.
- Gibb, F. & Matthaiakis, I. (2007). A framework for assessing web site localisation. *The Electronic Library* 25 (6), 664-678.
- Gläser, R. (1990). *Fachtextsorten im Englischen*. Tübinga: Narr.
- Göpferich, S. (1995). *Textsorten in Naturwissenschaften und Technik. Pragmatische Typologie-Kontrastierung-Translation*. Tübinga: Gunter Narr.

- Harper, M., O'Connor, K. & Simpson, M. (1999). *Quality Assessment: Fitting the Pieces Together*. Toronto: OSSTF.
- Holz-Mänttari, J. (1984). *Translatorisches Handeln. Theorie und Methode*. Helsinki: Academia Scientiarum Fennica.
- Hönig, Hans G. (1998). Positions, Power and Practice: Functionalist Approaches and Translation Quality Assessment. In C. Schäffner (Ed.), *Translation and Quality* (pp.6-34). Clevedon: Multilingual Matters.
- House, J. (1977). *A Model for Translation Quality Assessment*. Tübingen: TBL Verlag Gunter Narr.
- House, J. (1997). *Translation Quality Assessment: A Model Revisited*. Tübingen: Gunter Narr.
- House, J. (2001). Translation Quality Assessment: Linguistic Description versus Social Evaluation. *Meta* 46 (2), 243-257.
- Hurtado Albir, A. (2001). *Traducción y Traductología. Introducción a la traductología*. Madrid: Cátedra.
- Jeney, C. (2007). *Writing for the Web: a Practical Guide*. Columbus, OH: Pearson Prentice Hall.
- Jiménez-Crespo, M. A. (2009). Conventions in localisation: a corpus study of original vs. translated web texts. *Jostrans: The Journal of Specialized Translation*, 12, 79-102.
- Jiménez-Crespo, M. A. (2008a). *El proceso de localización web: estudio contrastivo de un corpus comparable del género sitio web corporativo*. Unpublished doctoral dissertation, University of Granada, Spain.
- Jiménez-Crespo, M.A. (2008b). Caracterización del género 'sitio web corporativo' español: análisis descriptivo con fines traductológicos. In M. Fernández Sánchez & R. Muñoz

Miguel A. Jiménez-Crespo

- Martín (Eds.), *Aproximaciones cognitivas al estudio de la traducción e interpretación* (pp. 259-300). Granada: Comares.
- Kenny, D. (2001). *Lexis and Creativity in Translation. A corpus-based study*. Manchester: St. Jerome.
- Kussmaul, P. (1995). *Training the Translator*. Amsterdam-Philadelphia: John Benjamins.
- Larose, R. (1998). Méthodologie de l'évaluation des traductions. *Meta* 43 (2): 163-186.
- Laviosa, S. (2002). *Corpus-based Translation Studies*. Amsterdam: Rodopi.
- LISA. (2003). *Localization Industry Primer*. D. Fry (Ed.) The Localization Industry Standards Association (LISA). Retrieved from <http://www.lisa.org/products/primer.html>.
- LISA. (2004). *Localization Industry Primer*, 2nd Edition. Geneva: The Localization Industry Standards Association (LISA).
- Lommel, A. (Ed.) (2007). *LISA Globalization Industry Primer*. Romainmôtier, Switzerland: LISA.
- López Rodríguez, C. I. & Tercedor Sánchez, M. (2008). Corpora and Students' Autonomy in Scientific and Technical Translation training. *Jostrans, Journal of Specialized Translation* 9, 2-19.
- Martínez Melis, N. & Hurtado Albir, A. (2001). Assessment in Translation Studies: Research Needs. *Meta* 47 (2), 272-287.
- Melby, A. K. (2006). MT+TM+QA: The Future is Ours. *Tradumatica* 4. Retrieved from <http://ddd.uab.cat/pub/tradumatica/15787559n4a4.pdf>
- Neubert, A. & Shreve, M. (1992). *Translation as Text*. Kent, Ohio: Kent State University.

- Nielsen, J. & Tahir, M. (2002). *Homepage usability: 50 Websites deconstructed*. Indianapolis: News Riders.
- Nobs, M. (2006). *La traducción de folletos turísticos: ¿Qué calidad demandan los turistas?*. Granada: Comares.
- Nord, C. (1997). *Translating as a Purposeful Activity. Functionalist Approaches Explained*. Manchester: St. Jerome.
- Nord, C. (1996). El error en la traducción: categorías y evaluación. In A. Hurtado (eEd.), *La enseñanza de la traducción, col. Estudis sobre la traducció 3*, (pp. 91-107). Castelló: Universitat Jaume I.
- Nord, C. (1991). *Text Analysis in Translation: Theory, Methodolgy and Didactic Application of a Model for Translation-Oriented Text Analysis*. Amesterdam/Atlanta: Rodopi.
- O'Hagan, M. (2009, April). *An exploration of the crowd as assessors of translation: translator training of popular entertainment in the context of crowdsourcing*. Paper presented at Monterey Forum 2009 Principles and Practices of Assessment in the Training of Translators and Interpreters, Monterey, CA.
- O'Hagan, M. & Ashworth, D. (2003). *Translation-Mediated Communication in a digital World: facing the challenges of Globalization and Localization*. Clevendon: Multilingual Matters.
- Olohan, M. (2004). *Introducing Corpora in Translation Studies*. London: Routledge.
- Ørsted, J. (2001). Quality and efficiency: Incompatible elements in Translation Practice? *Meta* 46, 438-447.

Miguel A. Jiménez-Crespo

Pierini, P. (2006). Quality in Web Translation: An Investigation into UK and Italian Tourism Web Sites. *Jostrans, The Journal of Specialized Translation* 8, 85-103.

Pym, A. (2003). What localization models can learn from Translation Theory. *The LISA Newsletter: Globalization Insider* 12 (2).

Pym, A. (2004). *The Moving Text: Localization, Translation and Distribution*. Amsterdam-Philadelphia: John Benjamins.

Reiss, K. & Vermeer, H. J. (1984). *Grundlegung einer Allgemeinen Translationstheorie*. Tübingen: Niemeyer.

Rothe-Neves, R. (2002). Translation Quality Assessment for Research Purposes: An Empirical Approach. *Cadernos de Tradução: O processo de Tradução* 2 (10), 113-131.

Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Robbins, S. S. & Stylianou, A. C. (2003). Global corporate web sites: an empirical investigation of content and design. *Information & Management* 40, 205-212.

Sager, J. (1989). Quality and Standards: The Evaluation of Translations. In C. Picken (Ed.), *The Translator's Handbook* (pp. 91-102). London: ASLIB.

Shreve, G. M. (2006). Corpus Enhancement and localization. In K. Dunne (Ed.), *Perspectives on Localization* (pp. 309-331). Amsterdam-Philadelphia: John Benjamins.

Stubbs, M. (1996). *Text and Corpus Analysis: Computer-assisted studies of language and culture*. London: Blackwell Publishers.

Toury, G. (1995). *Descriptive Translation Studies and beyond*. Amsterdam-Philadelphia: John Benjamins.

Pragmatic and Functionalist Aspects in Localization

Waddington, C. (2001). Different Methods of Evaluating Student Translation: The Question of Validity. *Meta* 46, 312-325.

Williams, M. (2003). *Translation Quality Assessment*. Ottawa: Ottawa University Press.

Wright, S. E. (2006). Language Industry Standards. In K. Dunne (Ed.) *Perspectives on Localization* (pp. 241-278). Amsterdam-Philadelphia: John Benjamins.

Evolution of User-generated Translation: Fansubs, Translation Hacking and Crowdsourcing

Minako O'Hagan

School of Applied Language and Intercultural Studies, Dublin City University

Abstract

Most conspicuous initially with Japanese anime fansubs, fan-based translation has been developing since the 1980s. In the advent of widespread availability of Web 2.0 platforms, fan translation has become a global phenomenon despite its dubious legal status. Extending the emerging interest in fansubs and scanlation in translation studies to the little discussed translation hacking by video game fans, this article brings readers' attention to participatory culture manifest in user-generated content in the field of translation and localisation. The article describes the evolution from unsolicited fan translation to solicited community translation now called crowdsourcing and considers them in the framework of user-generated translation (UGT). The article provides interdisciplinary perspectives, drawing on insights from media and game studies to address UGT which could profoundly impact the profession of translation and localisation as new technological environments unleash the technical competence, genre-knowledge and unparalleled devotion of the otherwise untrained Internet crowd as translators.

Introduction

The birth of localisation in the early 1980s signalled the beginning of a new era of translation in the transition to the digital age (O'Hagan & Ashworth, 2002). The requirement to translate digital content such as computer software had a profound impact on translation. Translators still translated words, but translation now needed to be formulated and coded in a specific form suitable for a given digital platform, in turn demanding new techniques and processes of localisation. The term "localisation" was introduced to signify the extra dimensions required to incorporate the translator's work into digital products. This also meant new ways of working with Computer-aided Translation (CAT) tools which have now seeped through to other areas of more conventional translation. While localisation has become a well established industry today, the conceptual boundary between translation and localisation has remained unclear and theorisation of localisation is still ongoing in the field of Translation Studies (TS) (Pym, 2009). For this reason, the term "Translation" with a capital T is used in this article to encompass both translation and localisation where appropriate. Such an ambiguity also reflects the fluid nature of the domain of localisation which continues to expand its boundary in the tide of globalisation of an increasing range of digital media. For example, video game localisation has added new dimensions to utility software localisation. Modern video games are interactive multimedia systems, with their rich content comprising written text, graphics, cut-scenes (movies), sounds, etc realised in a highly complex technological system. Game localisation is an emerging area of study within TS (see Bernal, 2006; Mangiron, 2007; O'Hagan, 2006) and the topic of fan translation of games is even less explored. Game localisation involves techniques used for software localisation and also elements of audiovisual translation where translated

words in different modes are integrated into the electronic medium. It belongs to a highly specialised area of translation, requiring special training designed for interactive multimedia applications for entertainment. Given the specialised nature of the tasks involved, it is doubly surprising that some gamers tackle the challenge without formal training or provision of technical support. They seem to acquire the necessary knowledge and skills through online collaboration with like-minded gamers and knowledge-bases created by the game community in addition to their own game playing experience and often extensive research.

Following the recent deployment especially in localisation of user-based crowdsourcing model, this article discusses the phenomenon of user participation in otherwise highly specialised areas of professional Translation practice. Media and cultural studies scholars such as Henry Jenkins (2006) see the trends of various media users extending their involvement in media to a much more active and deeper level to form a “participatory culture” facilitated by new media technologies. Progressing from the first generation Web, Web 2.0 ushered in the age of users as producers who freely contribute a wide range of content for public consumption as exemplified in *Wikipedia* and *YouTube*. In this context, the concept “user-generated content” (UGC) has become highly resonant with today’s Internet which itself is considered as user-generated at one level (Flew, 2008: 35). Without the Internet’s extensive and ubiquitous coverage albeit still excluding many developing countries, UGC would never have reached the level of significance we are witnessing today. It was indicative that *Time* chose “you” the user as the Person of the Year in 2006 (Grossman, 2006), marking the special and unprecedented role played by users. *Time*’s cover read: “Yes, you. You control the Information Age” (Time, 2006, December 13). With the focus on users in today’s media landscape, this article examines Translation which is

Minako O'Hagan

also coming under the increasing control of users as global electronic communications leverage the collective intelligence of users or the Internet crowd in a completely new way, forming, in the best case scenario, “the wisdom of crowds” (Surowiecki, 2004).

For the purpose of this article the meaning of the key term “user” needs to be clarified in relation to the ideas behind UGC. From a media studies perspective, Flew (2008, pp. 35-36) explains UGC as “the way in which users as both *remediators* [original emphasis] and direct producers of new media content engage in new forms of large-scale participation in the digital media spaces”. Drawing on this description, I use the term “user-generated translation” (UGT) in this article to mean a wide range of Translation, carried out based on free user participation in digital media spaces where Translation is undertaken by unspecified self-selected individuals. The user in UGT therefore is somebody who voluntarily acts as a “remediator” of linguistically inaccessible products and “direct producer” of Translation on the basis of their knowledge of the given language as well as that of particular media content or genre, spurred by their substantial interest in the topic. In the context of game localisation performed by users, the term “gamer” is used to mean users/players of games involved in Translation as it also connotes committed game players who tend to have amassed the necessary knowledge to undertake the often technically intricate task of game localisation. DePalma & Kelly (2008) use the term CT³ in reference to “community translation”, “collaborative technology and processes” and “crowdsourcing” which are all merging to create “translation that is generated of, by, and for the people”. Such characterisation underpins the central concept behind UGT aligned with social networking – a buzz word publicised by new communications platforms such as *Facebook*, *Second Life* and *Twitter*. Implicit in social networking platforms is their recognition of the inherent human nature

to socialise and interact, which will also facilitate tasks to be performed in different professional areas, including Translation.

UGT provided in a voluntary manner without monetary reward raises issues relating to quality, ethics and the very survival of the Translation profession. Drawing on insights from media and game studies to understand the implications of technological developments affording UGT and also taking a sociological approach to shed light into the nature of social interactions among gamers in the game community, this article attempts to make sense of UGT from interdisciplinary angles. Taking the cases of a number of prominent fan translation practices to date and linking them to the recent crowdsourcing model, the article highlights the significant changes that could potentially transform the whole profession of Translation.

Developments of User-generated Translation: Fan Translation

In order to provide a historical perspective it is relevant to discuss earlier forms of UGT of note and how they have developed. In the advent of mass-scale connectivity on web-based platforms, like-minded people are able to congregate and work together on self-selected projects forming what Rheingold (1993) had earlier called the “virtual community”. In the wake of the new generation of Internet with Web 2.0, members of the virtual community also came to gain access to a better and wider range of tools required for producing various content in electronic form while collaborating with virtual community members. In this environment, Bey et al. (2006) note that two types of volunteer translator networks arose: (1) mission-oriented and (2) subject-oriented. The former refers to highly-coordinated groups mainly engaged in technical translation of documents related to open source software products whereas the latter refers to translations of online documents performed by individuals chiefly to advocate certain points of

Minako O'Hagan

view such as support of humanitarian causes. Both these examples seem to be motivated by the general spirit of free sharing of material by like-minded people for which self-appointed volunteer translators are willing to lend hands, be it further distributing of open source programs or pursuing various topics which strongly concern them. An argument often put forward in support of such volunteer networks especially for translation within technical domains is an alleged lack of sufficient subject knowledge on the part of professional translators to be able to translate specialised documents such as those related to open source software. The issue of domain-knowledge is a significant factor behind the development of UGT and also relates to fan translation networks especially conspicuous in popular culture genres, which can be considered as a third category of volunteer translators (O'Hagan, 2009).

Among the early adopters of new collaborating possibilities on the Internet have been various fan groups. The availability of the Internet and collaboration tools gave such fan groups means to express a collective voice irrespective of their physical location allowing their presence to become much more visible, and leading to the formation of global “fan cultures” (Hills, 2002). Furthermore the nature of fandom changed from that of passive spectators to active “prosumers” who are producers as well as consumers of products (Tapscott & Williams, 2006). Fan translation fits well in the category of prosumers where potential consumers of translations double as translation producers. The most well-established early form of UGT was fan translation of Japanese animation now widely known as anime. Fan-produced subtitles for anime, called fansubs, are circulated worldwide in media spaces. This UGT is produced chiefly by fans for fans (Díaz-Cintas & Muños Sánchez, 2006). Despite its relative long history the phenomenon of fansubs has only recently begun to be recognised in the field of TS (Munday,

2008, p.190) mainly by scholars working in the field of audiovisual translation (e.g. Díaz-Cintas & Muños Sánchez, 2006; Pérez González, 2006). Fansubs' tendency of not conforming to norms of audiovisual translation and their often experimental nature gave rise to the concept of "abusive subtitles" (Nornes, 1999/2004, 2007) in reference to their translation strategy which, Nornes claims, directs the audience to the original by way of challenging some of the conventional constraints imposed on subtitles.

Such norm-defying translation strategies seem to stem from fans' search for "authentic text" (Cubbison, 2005). The approach is closely linked to the original motivation behind the development of fansubs which emerged as a protest against the official often over-edited versions of anime typically aired in dubbed form on television networks outside Japan. The early appearance of fansubs goes back to the 1980s with the use of VHS tapes which imposed cumbersome and expensive processes of private subtitling. In contrast today's far more efficient and cheaper fansub productions are sometimes called "digisubs" as the latter mainly rely on digital technology (Leonard, 2005, p.11). Digisub practices leverage technological environments, ranging from file sharing in order to share raw content for translation and to distribute the finished product to chat rooms for discussion as well as multimedia authoring and subtitling tools. There is also a more recent phenomenon of fan translation referred to as "scanlation" of mainly Japanese comics called *manga* (O'Hagan, 2008). The term derives from "scanning" of raw material comprising printed pages of original *manga* and "translating" them. Similar to digisubs the whole process takes advantage of digital production and communication environments. Both digisubs and scanlation today seem to be mainly motivated by a desire to fill in the gap or delay in official translations since the worldwide recognition of anime in

Minako O'Hagan

particular seems to have moderated some of the most radical editing strategies of the early days. Both forms of UGT voluntarily undertaken by fans leverage global connectivity and online collaboration to produce and deliver translations and make them available in media spaces. As well noted by observers of fan translation (Cubbison, 2005; Pérez González, 2006), one of the defining characteristics is fans' savvy use of technology both to produce UGT as well as to manage their workflow shared among different team members who are normally located in different physical sites. While fans' non-conformity and their very existence may be frowned upon by the professional translator community, fan translation approaches seem to be gradually seeping through to commercial productions in the area of anime subtitles (Caffrey, 2009), including cases where well-reputed fan translation groups are sometimes hired to produce an official version. The quality of fan translation can be variable (Díaz-Cintas & Muños Sánchez, 2006), but there are instances of fans' apparent lack of formal translator training being compensated for by their genre-knowledge (O'Hagan, 2008).

Drawing on prior study of fansubs and scanlation as fan translation networks, key aspects of fan translation relevant to the present discussion can be summarised as: (1) collaboration afforded by technology allows fans to form purposeful task groups to undertake a project and achieve a production often in a manner comparable to professional production in terms of the workflow process and the timeframe¹; (2) despite the dubious legal status copyright holders have largely condoned the practice, partly due to the fact that fan productions can facilitate the exposure of the given product to a wider target language public, thus in effect providing considerable free pre-publicity; (3) fan translators willingly undertake a translation project

¹For a fan translation of an episode of TV anime the timeframe is usually set within one or two days of its original broadcast in Japan whereas for a chapter of manga in scanlation the deadline is set within the same day of release (O'Hagan, 2008, p.165).

without remuneration, indicating a strong motivational factor and (4) domain-knowledge possessed by fans may in some cases compensate for the lack of formal translator training .

These key characteristics will be revisited in the final section on crowdsourcing.

The Video Game Fan Community and UGC

This section first looks into wider activities within the video game fan community that are relevant to the current discussion on UGC. The video game community can be seen as a prime example of a UGC contributor, demonstrating the gamers' devotion and creativity, their subject and technical knowledge and the level of collective intelligence gathered as a group. The depth and breadth of fan activities leading to UGC in the game field have reached the level of a metaculture where "players construct elaborate out-of-game meeting places" as they "bring a game beyond screen" (Nielsen, S.E., Smith, J.H. & Tosca, S.P., 2008, p.157). One type of UGC within the game community relates to a practice broadly known as "poaching" where "fans creatively re-use content from other media" (ibid, p.158), creating so called "fan art or fan work". The term poaching itself suggests these fan activities often border on the legal breach of intellectual property and raise copyright issues. Opinions are divided even among fans and certain poaching activities are "not always supported by their fan community" (Burn, 2006, p. 102) even though the game industry has embraced rather invasive fan activities such as "modding" as discussed later. This is also something in common with anime fansubs where ethical issues are addressed differently from group to group and often condoned by the industry (Leonard, 2005). Despite the questionable nature of some fan activities, their participatory level seems to be on the rise judging from the increasing volume and visibility of fan produced materials available on the Internet. This is possibly due to benefits of fan activities perceived by

Minako O'Hagan

IP holders in the game industry. One pragmatic example is a “walkthrough” which provides a detailed step by step player guide in a manner of enactment of a particular game. Walkthroughs are created based on the contributor’s experience of the given game, showing a certain mastery of the game, which may or may not be endorsed by other players. They transfer a multimodal game to a form of procedural text typically written in imperative mood devoid of representational aspects of the game (Burn, 2006, p. 90). Burn further observes that these walkthroughs may read like a dispassionate factual text, but, in fact, are full of the passion of the authors (ibid, p. 91) who are driven to serve their fellow fans by helping them complete the game apart from some gamers’ exhibitionist tendency for the demonstration of expertise.

Walkthroughs can be considered to function as a “help file”, albeit unauthorised, for less experienced gamers. In this way, while legally dubious, some of the fan activities could be actually beneficial rather than detrimental to game publishers. The game industry generally seems to acknowledge such positive dimensions of active fan participations.

On further scrutiny into the video game fan activity of poaching, Burn (2006, p. 88) argues how such fan work can both “revere the original text, seeking to remain as true to it as possible” and at the same time “dramatically alter the original text, adapting it to express the particular interest of the fan or fan group”. These two opposing approaches are useful in contextualising the nature of fan activities in the video game domain and relevant to discussing game fan translation practices in particular. One of the prevalent forms of poaching is fan fiction commonly known as *fanfic* which is developed based on a range of media texts, including video game narratives. *Fanfic* has also extended into the audiovisual sphere where fans produce animated films to re-create elements of the original game in terms of story, characters, settings,

etc usually “for humorous or ironic purpose” (Nielsen et al., 2008, p.159). This type of fan work may be developed as a mockery or parody of the original. One of the best known examples is redistribution of blatant translation errors typically attributed to earlier Japanese video game localisation. For example, the translated phrase in the game *Zero Wing* (1991) “all your base are belong to us” has become so widely circulated on the Internet via various forms of fan work that it has reached legendary status with the game community. There are many more similar examples where translation errors have brought the game to wider public attention through *fanfic* movies, texts and online discussion sites. By comparison, some other *fanfic* activities reveal their concerns to ensure that their fictitious renditions remain consistent with hard facts based on the original “canon” (Newman, 2008, p.59). While *fanfic* entails a creative extension arising from the original game, there are tendencies to follow the authentic information even by distinguishing the canon from non-canon versions among the same game series based on the involvement or the absence of the original development team. Fan message board discussions often reveal the considerable research the fan community is prepared to undertake and how fans construct legitimate arguments based on their research. Even official information may not be taken at face value and fans may insist on referring back to the original game source when arguing about localised games. Newman (2008) provides an illustrative example where such debates involved discussions on the original versus localised versions with Sega’s popular game title *Sonic the Hedgehog* originally released in 1990.

Elsewhere in the analysis of the same issue, fans return to what they consider to be trustworthy ‘original’ (that is, the Japanese manuals that accompany the canonical games that are provided in translation by one of the community who lends his intensive knowledge to the

Minako O'Hagan

group). The authority of these materials supersedes even the official Sega of America (SoA) and Sega Europe materials which are discarded as disloyal as they frequently invent new narratives and explanations rather than faithfully translating and duplicating the canonical Japanese materials (Newman, 2008, p. 61).

This stance is sometimes reflected in the fan translation approach which seeks to replicate the original voice similar to fansubs. This, in turn, highlights the different focus fans may have from the goal of commercial game localisation. The latter puts priority on adapting the game to the target market requirements so that the product is received as if it is a locally produced game, not a translation (Mangiron & O'Hagan, 2006). Newman observes (2008, p.65) fans' collective intelligence about certain game titles to have reached the extent that they are able to discern "those that are genuine inconsistencies in the authorially authentic narrative and those that are the products of poor or insensitive translation". These examples provide relevant background to understanding some fan perspectives on officially localised games. Fans' knowledge on the topic supported by their extensive research can be contrasted to the diminishing timeframe and lack of context under which professional translators and localisers are often forced to work in the field of game localisation. Furthermore, the situation can be frustrated by the fact that some official European language versions of Japanese games are derived from the games' English versions, leaving more room for loss of the original intent. These may have led to a degree of scepticism about some of the official versions by fans as they observe: "those involved with the official translations either take liberties with or are simply insufficiently well versed in the minute detail of the canon to produce a sensitive English language version". With some game series fans are more sensitive to "the lack of care in preserving the continuity and integrity of the

canon as envisioned by the originators” (Newman, 2008, p.61). Such critical attitudes of some fans arise from their knowledge and mastery of specific game titles, which must feed into how they undertake translation and localisation. Similarly these observations serve to explain how, in some cases, fans are compelled to undertake game localisation as part of fandom and extended participatory engagement with the medium. The breadth and level of engagement and collective knowledge accumulated by the video game fan community is apparent in the vast range of online activities and resources such as discussions in message boards and fan work creations. UGT by these fans often entails a depth of commitment on the part of the contributor which commercial Translation is rarely able to match. This seems to have resulted in the trust of fans being placed more on translations that “emerge from within the trusted community of gamers rather than those of regional marketing departments” (Newman, 2008, p. 64).

Of various activities by video game fans, a practice which best highlights the blurring boundary between the producer and the consumer in terms of UGC is the phenomenon known as “modding” in reference to modification of games. A “mod” ranges from a light cosmetic touch to a complete overhaul of the original game and is a further testimony to the extensive technical knowledge acquired by some fans. Interestingly mods are not only condoned but are endorsed by the game industry as a way to enhance the appeal of the original game and gather wider gamer community attention and extended engagement in the game. This is accommodated to the extent that some game developers officially make available development kits to facilitate the modding process as well as designing original games in such a way as to avoid any harmful impact of modding on the original game code by separating the main program from media files (Nielsen et al., 2008, p.161). The best known and by far the most successful example to date is

Minako O'Hagan

Counter-Strike which appeared in 1999 as a mod of *First Person Shooter* (FPS) game *Half-Life*.

The way modding has been harnessed by the game industry is by controlling it under official agreement of EULA (End User License Agreement) which curtails any profit gained on the part of the modifiers from their activities. Modding exemplifies how the game industry dealt with UGC whereby user expertise and creativity are encouraged to the advantage of the industry. The next section homes in on game fan activities directly in the area of UGT.

Fan Translation of Video Games: Translation hacking

Unlike business software development which is centred mainly in the USA and thus originates in English, one of the major game developing countries is Japan, adding Japanese to main source languages of game localisation alongside English. There is a close link between anime fansubs and fan translation of video games, both sharing Japanese as a major source language where the initial goal was to make Japanese exclusive materials accessible especially to the English-speaking world. Today fan translation is performed into many different languages. Newman (2008, p.156) notes a limited availability of localised versions in the early days of Japanese Role Playing Games (RPGs) due to their especially “copious quantities of text”, signifying a considerable investment for game publishers. With the tendency of RPG titles to pack in game play times which may stretch to over 100 hours, the volume of game content subject to localisation is increasing. The decision on whether to localise thus has greater cost implications, and means that some territories are left out in the cold with some games unlocalised. Faced with the lack of localisation typically, but not exclusively, of Japanese RPG titles, some dedicated fans have come to organise their own translation, referred to as “translation hacking” which is a type of “ROM-hacking”. It entails modifying the binary ROM (read-only

memory) image of the game. Translation hacking involves a hacker and a translator working together to extract the relevant text from the ROM and to replace it with a translated script.

While the motivation and the level of enthusiasm can be comparable to those found in fansubs and scanlation activities, the fan undertaking of video game localisation entails significantly more time-consuming and technically challenging tasks (Newman, 2008, p. 158). It is this technical knowledge and intensive interest in the game development process by some gamers that characterise the unique nature of fan translation of games.

To further explain the technical background to translation hacking, it is relevant to point out the distinction between PC games and those played on dedicated consoles such as *PlayStation*. Games played on PCs have been more amenable to making changes as computer programs more readily allow external modifications, for example, with patches to update or debug the programs. By comparison, console games represent a tightly sealed closed system, making external changes more difficult. For this reason while PC games have been subject for a long time to user modifications, this is a newer development with console games (Taylor, 2007, p.234). Translation hacking for console games is closely linked to the availability in the late 1990s of console emulator software which turns a PC into a virtual game console. ROM hacking, as explained by the practitioner (Mandelin cited in Parkin, 2008), involves a hacker locating the game's font to produce what is known as a table, which in turn allows text data to be identified and then to be copied ("dumped") in a file for translation. The translation is released as a ROM patch applied to the original ROM to effect the translation and is played on a console emulator. This process is highly technical, sometimes involving such tasks as changing the original (Japanese) font to include English letters and for large projects the hacker may even

Minako O'Hagan

create a special program to make the translators' tasks easier (Mandelin cited in Parkin, 2008).

In addition to the increasing volume of text to be translated, the complex game structure provides further challenge even to identify all elements which need to be localised.

The most recent high-profile fan translation example is that of Japanese RPG title *Mother 3* (2006) originally released for *Nintendo Game Boy Advance*. The English fan translation was made available in October 2008 after nearly two years' worth of work. The series has had a serious following by overseas fans who produced a copious amount of *fanfics* and *fanart* and yet the game had not been localised probably because of poor sales of *Mother 2* in the US (Newman, 2008, p.157). The scale of popularity of the fan translated version can be gauged by the number of translation patch downloads exceeding 100,000 during its first week of release (Parkin, 2008).

In an interview by the game website *Gamasutra*, the main fan translator involved in the project Clyde Mandelin, a professional translator working in the game localisation industry, responds to the question regarding the legal issue by stating: "we were fully prepared and willing to stop our work immediately if there had been any word from the game's IP owner" (Parkin, 2008, p.2).

True to his words, the project's official website "Mother 3 Fan Translation"

(<http://mother3.fobby.net/faqs/>) has the following declaration:

As we've mentioned throughout the history of the project, our team has agreed that this translation project will come to a full stop if/when we hear that *Nintendo* has chosen to pursue an official translation of the game. Our only goal is to get MOTHER 3 in the hands of the fans.

Mandelin stresses how fan translation is a "hobby" undertaken as "out of love" and generally there is a mutual understanding of fans' intention between fan translators and the IP owners

(Parkin, 2008, p. 2). Feedback from fans on the fan translated version seems to assume such an amicable relationship whereby many of them openly voice their hope that *Nintendo* will notice such fan endeavours and be encouraged to localise their games (see *Siliconera Speaks Up: The Mother 3 Fan-Translation*, 2009). While the profile of a fan translator is by definition not a trained professional, Mandelin is an example of a fan translator turned professional yet still retaining his former role. Mandelin maintains that the experience gained from fan translation has been invaluable to his profession and adds: "every pro translator I know is/was a fan translator" (Parkin, 2008, p.2). This points to a close linkage between fandom and professional translation in the field of game localisation and implies the benefit of working as part of fan translation communities. The latter seem to be able to provide, if by accident, a situated authentic learning environment, likely boosting the process of preparing trainees for the game localisation industry. As such fan translation networks can be seen as an unlikely but perfect example of a social constructivist approach applied in translator training outside the formal translation school" (O'Hagan, 2008). The phenomenon of fan translation illustrates how technological environments have provided a solution, albeit legally questionable, to the lack of availability of official translation allowing user-translators to step in and showcase their domain-specific knowledge and technical skills. In this way today's enhanced Web 2.0 environments are harnessing UGC and UGT. What is particularly new, however, is the trend where these mostly unsolicited and often legally illicit user-based translation activities are now turning into solicited legitimate activities developing as an increasingly accepted business model.

Crowdsourcing: Legitimatising UGT

Most forms of UGT discussed above stand on legally shaky ground regardless of whether their original copyright holders are likely to take legal action. Leonard (2005) argues that the dismissive attitudes towards fansubs by Japanese copyright holders of anime in the early days have, in fact, served them well by benefitting them in the long run by facilitating a “demand formation” of anime at the time of relative obscurity of the genre in the US and elsewhere. Similar to fan translation of games, many fansub groups profess to work on the principle that when an officially translated version comes out in the market, they will retrieve their fan translated version although in reality such actions may make little difference once the materials have been circulated online. Digitisation of products and the availability of ubiquitous broadband communications networks have created an increasing difficulty in quashing digital content piracy notably suffered by the film and music industries, leading to closing of some file sharing sites. The popularity of UGC and now the rise of UGT appear to be adding to the continuous problem which arises from the very nature of going digital where users control information as observed by *Time* (Grossman, 2006). The term “translation hacking” used to describe fan translation of video games is also indicative of the breach of legality relating to computer hackers. Nevertheless the instances of UGT discussed so far show how powerfully motivated users will engage in laborious time-consuming tasks in the spirit of sharing, which, in turn, is more or less accepted in the otherwise fiercely commercially-oriented, litigation-bound game industry.

A new business model which has recently come to be known as crowdsourcing takes a step further the trend of leveraging free labour of love. Howe (2008, p.6) explains his choice of

this naming in 2006, sensing the unprecedented impact of combined dynamics arising from UGC as something which is rapidly shaping our culture and economy. Taking advantage of its core-business expertise of social networking, *Facebook* launched crowdsourcing in 2007 to let their users translate selected parts of the *Facebook* website into different language versions. It provides one of the first full-fledged cases allowing us to observe how this model works in the field of web localisation. As of September 2008 *Facebook* had involved over 30,000 users around the world in their project of translating its website into 16 launched languages (Losse, 2008). *Facebook* uses the Internet crowd to translate selected strings which form part of their websites. By developing a user-friendly translation application, called *Facebook Translations*², *Facebook* leverages their users' enthusiasm and mass collaboration also to evaluate the contributed translations by a user voting mechanism (see Figure 1). The results of the voting in turn feed into translator leader boards, publicising and recognising the level of activity of top contributors both quantity- and quality-wise. This clearly provides the contributors with a motivational factor, showing the number of positive votes they received for their translation. While other unsolicited forms of UGT discussed above leave translation quality as an open question that other users may or may not comment on, this model directly elicits user comments on quality. According to *Facebook* their motivation for using a crowdsourcing model is primarily not a matter of cost saving. It points out that most savings which may have been made on translation in fact have been spent on developing its technology platform with the translation application. Furthermore they engage commercial translation service providers for translating

² At the time of writing *Facebook* is awaiting the outcome of its patent application for "techniques for translating text in a social network" filed in December 2008 with the US Patent and Trademark office. The full application is accessible at: <http://appft1.uspto.gov/netacgi/nph-Parser?Sect1=PTO1&Sect2=HITOFF&d=PG01&p=1&u=%2Fnetahtml%2FPTO%2Fsrchnum.html&r=1&f=G&l=50&s1=%2220090198487%22.PGNR.&OS=DN/20090198487&RS=DN/20090198487>

Minako O'Hagan

more than half of *Facebook* content in its supported languages and also for evaluation of translations done by the users, validating the users' own democratic method of quality control (DePalma & Kelly, 2008, p.15).

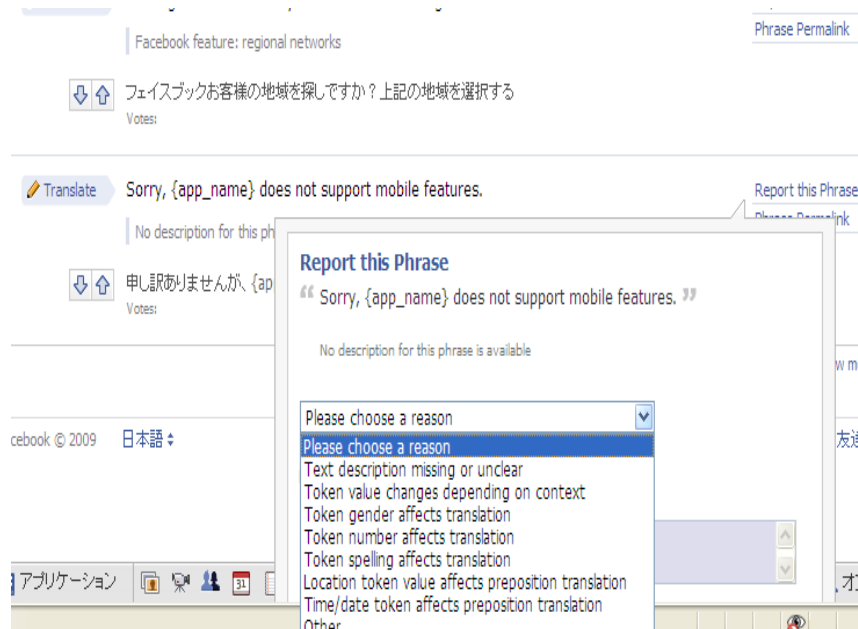


Figure 1: Translation Quality Feedback by User Voting on Japanese *Facebook* (source: www.facebook.com)

The crowdsourcing model as applied by *Facebook* confirms some of the prior findings on the characteristics of fan translation: the successful deployment of a well-motivated collaborative community for a purposeful task. *Facebook* leveraged the inherent interest of the site's users to make the social networking website accessible in a wide range of languages and also provided a mechanism to formally recognise their contributions using the leader board concept which in turn is supported by the crowd's opinion about the quality of contributed translations. Both these approaches seem to serve well to encourage further user participation either as translators or translation reviewers. Another aspect identified by prior study regarding users' domain-knowledge was also recognised by *Facebook* where the user group was reported to have

outperformed the professional translators with certain translations due to the former's familiarity with the inner workings of *Facebook* (Losse, 2008). This is presented as part of the argument by *Facebook* to legitimise the use of the crowdsourcing model whereby pointing out shortcomings of professional translators who had never used *Facebook* before (Losse, 2008).

This leaves us with one further issue regarding the ethical question of for-profit business ventures seemingly taking advantage of free labour of users. A recent case in point is the outrage expressed by its professional translator members when *LinkedIn*, a high-profile social networking site designed for professionals, in effect, asked how they would like to participate in crowdsourced translation of the *LinkedIn* website for free (Kelly, 2009). The translators' negative reaction was such that it led to the formation within *LinkedIn* of the group "Translators against Crowdsourcing for Commercial Business". Reporting this strong reaction Kelly (2009) argues that this is due, in part, to a lack of understanding by translators of what crowdsourcing entails as she argues that "it is simply another method of working in the digital age". Kelly further compares this new phenomenon of crowdsourcing to the emergence of computer-assisted translation (CAT) tools which are widely accepted and leveraged by the professional translation community today and yet may once have been considered as a threat by freelance translators who were not familiar with the technology. This case illustrates how UGT is likely to shake the profession further as more businesses become interested in trialling this model, especially if more, albeit variously qualified, users are prepared to participate in crowdsourced translation and start to undercut professional translators.

Conclusions

Based on observations of the evolving fan translation phenomenon including the highly technical translation hacking of video games and the emerging crowdsourcing model, this article argued that the concept of user-generated translation (UGT) is likely to have far-reaching impacts on the Translation profession. The digital revolution has come to empower mere “users” in an unprecedented manner whereby enabling them to contribute their genre-knowledge and to become part of a community of practice of a new kind. The technological advancements which are increasing the demand for Translation are now also masking the highly specialised nature of the work of translators and localisers. While professional language services are being sought by businesses to compete in global markets, users help translate *Facebook's* website into Spanish and German in just one week with its French version taking only 24 hours (DePalma & Kelly, 2008, p.12). Casual observers of crowdsourcing may have indeed reached the conclusion that translation and localisation are something freely and readily achievable by the Internet crowds where the boundary between the professional translator and the amateur is no longer clear. Today's rapidly transforming technological environments are challenging professionals in various fields. As the UGT phenomenon starts to enter into professional and academic discourses, we all need to engage in the debate with open mind and endeavour to understand the emerging contexts by establishing the facts so that we can critically assess the situation. There are unanswered questions of ethical issues and quality concerns which arise from the new model of crowdsourced translation. At the same time, there are also questions relating to genre-knowledge and general motivation of professional translators and localisers, which may stem from increasingly squeezed working conditions. The new digital world which favours openness,

Evolution of User-generated Translation

collaboration and sharing calls into question perhaps some of the hidden issues which are already plaguing the current work mode of professional translators and localisers. The new rules being established by the world of Web 2.0 are forcing the Translation profession to confront the new and the old issues. Perhaps a fresh insight may come from the process of reflection so that the unfolding changes can be leveraged to further enhance and progress the profession, and not bring about its demise. A clear way towards the decline will be to ignore the changes and insist on forever keeping the old way of working.

References

- Bernal, M.(2006). “On the Translation of Video Games”, *The Journal of Specialised Translation* 6: 22-36.
- Bey, Y., Boitet, C., & Kageura, K. (2006, May). The TRANSBey prototype: An online collaborative Wikibased CAT environment for volunteer translators. In E. Yuste (Ed.), *Proceedings of the 3rd International Workshop on Language Resources for Translation Work, Research & Training (LR4Trans-III)* (pp. 49–54). LREC 2006: Fifth International Conference on Language Resources and Evaluation. Magazzini del Cotone Conference Centre, Genoa, Italy. Paris: ELRA/ELDA (European Language Resources Association/European Language Resources Distribution Association).
- Burn, A. (2006). Reworking the text: online fandom. In Carr, D., Buckingham, D., Burn, A. and Schott, G. (eds), *Computer Games: Text, Narrative and Play*. Cambridge: Polity Press, pp. 72-87.
- Caffrey, C. (2009). *Relevant abuse? Investigating the effects of an abusive subtitling procedure on the perception of TV anime using eye tracker and questionnaire*. Unpublished PhD thesis, Dublin City University.
- Chandler, H. (2005) *The Game Localisation Handbook*. Massachusetts: Charles River Media.
- Cubbison, L. (2005). ‘Anime fans, DVDs, and the authentic text’, *The Velvet Light Trap*, 56, (1), 45-57.
- DePalma,D. A. & Kelly, N. (2008, December). *Translation of, for, and by the People: How User-Translated Content Projects Work in Real Life*. Massachusetts: Common Sense Advisory.

- Díaz-Cintas, J. and Muños Sánchez, P (2006), 'Fansubs: Audiovisual translation in an amateur environment', *The Journal of Specialised Translation*, 6, 37–52. Retrieved from: http://www.jostrans.org/issue06/art_diaz_munoz.php [Last accessed: May 2009].
- Flew, T. (2008). *New Media: An introduction*. 3rd edition, Oxford and New York: Oxford University Press.
- Grossman, L (2006, December 13). Times Person of the Year: You, *Time*.
- Hills, M. (2002). *Fan Cultures*. London, New York: Routledge.
- Howe, J. (2008). *Crowdsourcing: How the power of the crowd is driving the future of business*. London: Random House.
- Jenkins, H. (1992). *Textual Porchers*. New York: Routledge.
- Jenkins, H. (2006). *Fans, Bloggers, and Gamers: Exploring Participatory Culture*. New York and London: New York University Press.
- Kelly, N. (2009, June 19). Freelance Translators Clash with LinkedIn over Crowdsourced Translation. Retrieved from: <http://www.globalwatchtower.com/2009/06/19/linkedin-ct3/> [Last accessed: June 2009].
- Leonard, S. (2005). "Progress against the law: anime and fandom, with the key to the globalisation of culture", *International Journal of Cultural Studies*, 8 (3), 281-305.
- Losse, K. (2008). "Achieving Quality in a Crowd-sourced Translation Environment". Keynote Presentation at the 13th Localisation Research Conference Localisation4All, Marino Institute of Education, Dublin, 2-3 October, 2008.

Minako O'Hagan

Mangiron, C. (Ed.) (2007). *Revista Tradumàtica*. Núm. 5: "La localització de videojocs".

Retrieved from: <http://www.fti.uab.es/tradumatica/revista/num5/Sumari.htm> [Last accessed: May 2009].

Mangiron, C. & O'Hagan, M. (2006). Game Localisation: Unleashing Imagination with "Restricted Translation". In Diaz-Cintas, Orero, Remael (eds.) *the Journal of Specialised Translation*. Issue 06, 10-21. Retrieved from:

http://www.jostrans.org/issue06/art_ohagan.php#abstract [Last accessed: May 2009].

Munday, J. (2008). *Introducing Translation Studies: Theories and applications*. 2nd edition. New York: Routledge.

Newman, J. (2008). *Playing with Videogames*. Oxon and New York: Routledge.

Nielsen, S.E., Smith, J.H. and Tosca, S.P. (2008). *Understanding Video Games: The Essential Introduction*. New York and London: Routledge.

Nornes, Abé M. (1999/2004). 'For an abusive subtitling', in Venuti, L. (Ed.), *The Translation Studies Reader* (Second edition). New York/London: Routledge, pp. 447–469.

Nornes, Abé M. (2007). *Cinema Babel: Translating global cinema*, Minneapolis, London: University of Minnesota Press.

O'Hagan, M. (2006). "Manga, Anime and Video Games: Globalizing Japanese Cultural Production". *Perspectives: Studies in Translatology Special Issue on Manga, Anime and Videogames*, 14 (4), 243-247.

O'Hagan, M. (2008). Fan Translation Networks: An Accidental Translator Training Environment? In Kearns, J. (ed). *Translator and Interpreter Training: Methods and Debates*. London: Continuum, pp.158-183.

- O'Hagan, M. (2009). Teletranslation. In Pagani, M. (Ed.), the *Encyclopedia of Multimedia Technology and Networking*. 2nd edition, vol. III. Hershey, P.A.: Idea Group, pp 1379-1386.
- O'Hagan, M. & Ashworth, D. (2002). *Translation Mediated Communication in a Digital World: Facing the challenges of globalization and localization*. Clevedon: Multilingual Matters.
- Parkin, S. (2008, December 26). *You Say Tomato: A pro on fan-translating Nintendo's Mother 3*. Retrieved from:
http://www.gamasutra.com/view/feature/3891/you_say_tomato_a_pro_on_.php [Last accessed: June 2009].
- Pérez González, L(2006), 'Fansubbing anime: Insights into the "butterfly effect" of globalisation on audiovisual translation', *Perspectives*, 14 (4), 260–277.
- Pym, A. (2009). *Exploring Translation Theories*. New York: Routledge.
- Rheingold, H. (1993). *The Virtual Community: Homesteading on the Electronic Frontier*. Reading, Mass: Addison-Wesley.
- Siliconera Speaks Up: The Mother 3 Fan-Translation (2008, 16 November). Retrieved from:
<http://www.siliconera.com/2008/11/16/siliconera-speaks-up-the-mother-3-fan-translation>
[Last accessed: June 2009].
- Surowiecki, J.(2004) *The Wisdom of Crowds: Why the many are smarter than the few*. London: Abacus.
- Tapscott, D. & Williams, A. (2006). *Wikinomics: how mass collaboration changes everything*. New York: Portfolio.

Minako O'Hagan

Taylor, L. (2007). Platform Dependent: Console and Computer Culture. In Williams, J.P. &

Smith, J.H.(eds.), *The Players' Realm: Studies on the culture of video games and gaming*.

Jefferson, North Carolina and London: McFarland & Company, Inc, Publishers, pp. 223-

238.

Profile of a Terminologist in Localization Environments

Barbara Inge Karsch

Microsoft Corporation, Redmond, USA

Abstract

Terminology as a subject for research has been around for centuries. Even terminology-related tasks have been carried out for many years. The profession of terminologist in the localization industry has been emerging and taken on a more distinct profile, as localization is maturing as an industry. This paper defines the profile of a terminologist in the localization industry. First, it sets important definitions, and then describes terminology work as an essential ingredient in the localization process, and conceptual analysis, in turn, as the essence of successful terminology work. It then gives examples of companies that employ terminologists and concludes by juxtaposing seemingly contradictory characteristics and skills of multilingual terminologists in localization settings.

Barbara Inge Karsch

Multilingual technical communication has greatly increased in volume over the last two decades, most notably in the computer industry. Initially, software designed for the English-language markets was translated at least into languages, such as German and Japanese. But economic globalization, technical advances and saturation of existing markets encouraged software producers to open up new language markets. In recent years, the volume of translation from languages other than English into languages other than English has increased as well (Beninato, 2008). As a result, software companies looked for ways to improve output quality and speed of the localization process.

Standardized terminology is the backbone of a quality-driven translation process. In a recent survey of 194 translation professionals, 87 percent confirmed that “having a terminology management process clearly improves the productivity” and 82 percent said “the time taken to research terms was their biggest difficulty” (SDL, 2009). While smaller and one-time localization endeavors can make due with a simple terminology list or spreadsheet, companies with an extremely high translation volume and a large number of languages invest in and maintain sophisticated terminology management systems (TMSs). And while most of the localization work is carried out across the globe by freelance translators, these companies have at least a few trained terminologists on staff. They influence and document standardization decisions, research and author definitions, drive the term creation process and explicate conceptual systems. The resulting information is documented in electronic records of terminology databases where it is accessible to all users in real time and online.

Like a lexicographical entry in a paper dictionary, a terminological entry is only as good as the information it contains. Apart from being as up-to-date as possible and complete by itself, an

entry must also be in line with its conceptual system—whether that system is made explicit in the TMS or not—as the terminology database is only as sustainable as the ontology formed by these conceptual entries. Terminology databases fail for three main reasons. First, if too many duplicates remain undiscovered, they will confuse users and clutter the database. Second, if conceptual entries are not “clean”, for instance, if universals and particulars are mixed up, users will not be able to make clear connections between concepts and again will end up confused. Third, if concepts are not based in or checked against reality. If accurate and complete entries placed in their concept system are the mandate, terminologists must be capable of finding and documenting that information. This simple demand translates into the profile of highly skilled professionals.

Who can fill such a role? This paper attempts to define the profile of a terminologist. First, it sets some definitions, and then describes terminology work as an essential ingredient in the translation process, and ontology work, in turn, as the essence of successful terminology work. It then gives examples of companies that employ terminologists and concludes by juxtaposing seemingly contradictory characteristics and skills of multilingual terminologists in localization settings.

Definitions

When computer scientists, business managers, terminologists and philosophers discuss terminology management, ontology, and related subjects, a fair amount of miscommunication occurs because there is little common understanding of the terms and concepts of these adjacent fields. Ontology, and even terminology projects, can only succeed if project participants, such as

Barbara Inge Karsch

engineers, project managers, terminologists, and their sponsoring management, speak the same language.

The terminology of terminology is fairly well standardized, and progress has been made to define and agree on ontology terminology. For example, the *NeOn* consortium is actively defining ontology terminology for activities in ontology creation. The current glossary contains approximately 60 definitions and is available through the *NeOn* Web site (*NeOn* consortium 2008). Furthermore, ISO TC 37 created an ontology task force in 2007, which has as one of its goals the creation of an ontology of ontologies, including definitions (Nistrup Madsen, 2008). This paper draws on definitions set forth, for example, in ISO 1087 for the terminology field, as well as on definitions created and used by ontologists (e.g. Smith, 2006).

At the core of a terminology, but also an ontology system, is the “stuff” that is described, documented, named, and connected to other “stuff.” In terminology management, one unit of this “stuff” is referred to as a *concept*. This term is used throughout this paper in the same sense as Smith defines *entity*, that is, “anything which exists, including objects, processes, qualities and states” (Smith, 2006). As shown in Figure 1, in a terminology database concepts are represented by unique IDs. To facilitate communication, they are also represented by definitions.

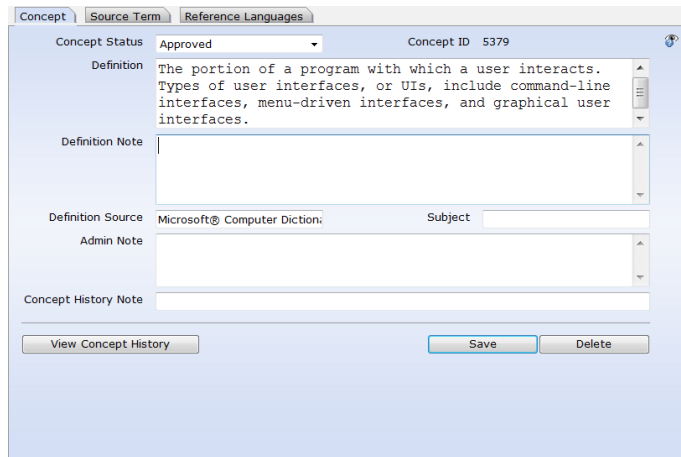


Figure 1. Conceptual entry for the term “user interface” in Microsoft Term Studio

Furthermore, concepts are represented by terms, appellations or symbols. In a terminology database, each such designator has its own terminological entry (“part of a terminological data collection [...] which contains the terminological data [...] related to one concept” (ISO 1087-1, 2000)). Terms representing the concept, such as the one described in Figure 1, in the same language are listed as synonyms with their own entries (see Figure 2), and terms representing the concept in another language are listed as equivalents.

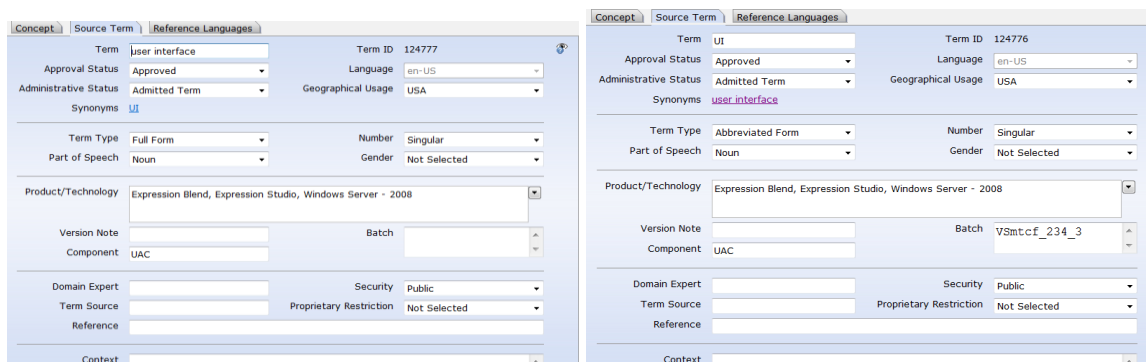


Figure 2. Term entries for synonyms representing the above concept in English

A visual representation of the terms *concept*, *taxonomy*, *concept system*, and *formalized ontology* may foster understanding. In Figure 3 illustration 1, the semiotic triangle (ISO 704,

2000) is used to represent a concept by itself. In the second illustration, several concepts are connected through generic relations; they form a taxonomy. Figure 3 shows concepts connected by multiple types of relations, for example, generic and partitive. These two types of relations are frequently the only ones used in TMSs. The difference between illustration 3 and 4 is that the relations may not be formalized or not even made explicit (here shown by the dashed line style). And yet they do exist between the concepts represented in a terminology database. Ideally, they are made explicit and even formalized, as indicated by the full lines in illustration 4, for example, to support the research for other languages, who will also attach their representations (e.g. in the form of terms) to the entry – here indicated by the arrows.

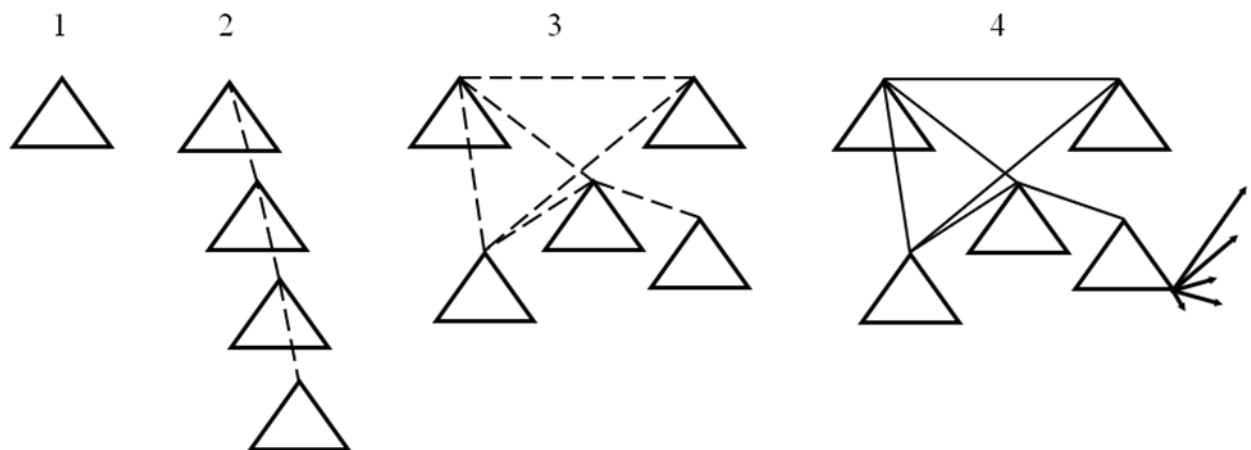


Figure 3. Visual representations of a concept (1), a concept hierarchy (taxonomy) (2), a concept system with implicit relations (3), and a concept system with explicit relations as well as multilingual representations (4)

Terminology Work: The Backbone of the Translation Process

Like most professions in the last three decades, the profession of a translator has changed significantly. Gone are the days of the typewriter and the need to live close to a library and one's customers. Today, ten localizers receive text chunks through e-mail, work in networked, virtual

environments that may even include subject-matter experts (SMEs), and hand back the finished files for compilation and publication as part of online help, a Web site, a printed document, or some other medium. While working methods have changed, consumers of localized software products have shifted their expectations, too. Sometimes they are satisfied to find information in their language at all, no matter the quality of the text. At minimum, when a functional or content-focused text, such as online help, is translated by a human being, the target language text must relay the information of the source text fully and free of errors (Reiß, 1986).

“Terminology is the primary means of communication and knowledge transfer between software developers and end users” (Schmitz, 2007). Terms and concepts must not only be used consistently, correctly, and unambiguously in the original document for the users of a source language product, users who buy translated products are also entitled to the same clarity and usability. To achieve clarity in the target language product and enable efficient use, translators must understand the ideas of a source text, the meaning of terms, and the relation between them. This is not always easy for a variety of reasons.

From the development of a new function to its representation in a target language, dozens of people may be involved. In a long chain of contributors, much of the original meaning of a concept may get lost by the time localizers see the first instance of a term in a text. If we take a prominent example such as the Ribbon, which was a new feature in all Microsoft® Office 2007 products, designers had an idea, developers implemented it, and content publishers wrote about it. They all had at least a vague notion of what the Ribbon was at the time they communicated about the concept. Microsoft localization project managers handed off the result of their communication, for example software files or user-assistance documents, to a vendor project

Barbara Inge Karsch

manager. They, in turn, distributed them to a number of freelance translators. A translator half-way around the world in their home office is then tasked with finding an appropriate term to express the meaning of Ribbon in a specific language. If the freelance translator does not see the implementation of the Ribbon (see Figure 4) or get an explanation of it—for example, in the form of an entry in a terminology database—the translator may not be able to find a better term than a literal translation or may decide to leave it in English. High-profile terms, such as Ribbon, generally receive enough attention by marketing in all relevant language markets that the meaning, as well as the preferred target name, is communicated even without a terminology database. But most technical terms do not receive this special and expensive attention, and without proper treatment in a terminology database they may end up being mistranslated.

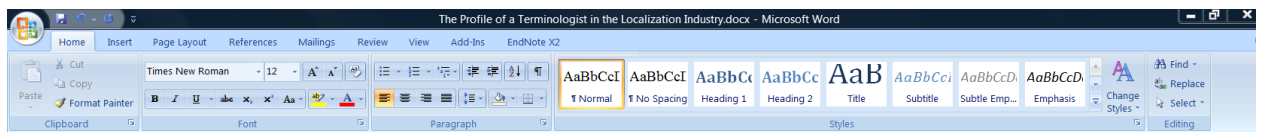


Figure 4. The Ribbon as displayed in Word 2007

In certain translation scenarios, translators may only get a glimpse into the world of the source text author. For example, if translation tools are poorly designed, translators may only see one string at a time. Figure 5 shows a screen print of the first translation tool used in the mid-90s at J.D. Edwards. The localizer is asked to translate the string “Swing” from English into German and arrived at the equivalent “Spätschicht,” which seems plausible for business software. But without context, the localizer either arrived at the translation through guess work or spent time to identify the meaning by extensive research. The fact that the Word Book Reference Language entry: “Auffahrt” in the upper portion of the screen is outright wrong certainly did not help.

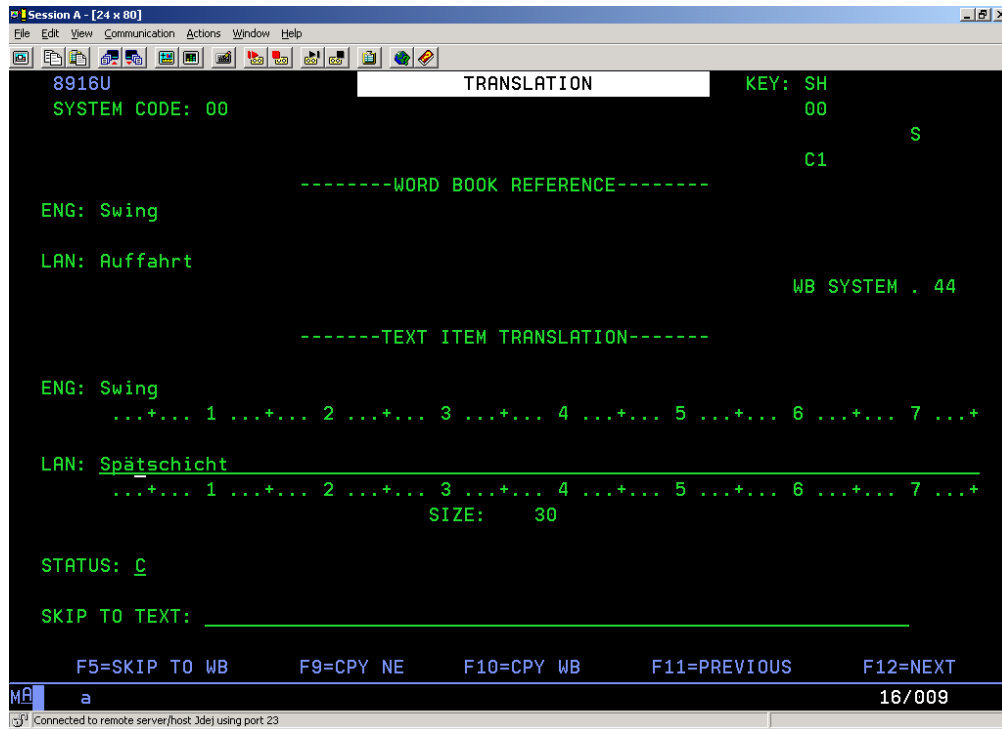


Figure 5. Translation software used at J.D. Edwards in the mid-90s

While a terminology database cannot eliminate additional, context-dependant research on the translators’ part 100 percent of the time, database entries can clarify the meaning of a term, provide contexts and indicate product information, and therefore greatly reduce the research time invested by translators. Furthermore, a terminology database can ensure that a team of several translators uses the same term for the same concept to produce a standardized text. The benefit of consistent terminology within a software product or between different interdependent products (such as Microsoft Office running on the Microsoft Windows® operating system) is easy to recognize.

Although a unit of translation (such as “Swing” in Figure 5) can be at the same time the term described and defined in a terminology database, translators and terminologists treat the

Barbara Inge Karsch

unit differently. A translator works in context and arrives at a target solution that is correct in that text. Eco says “in light of the full spectrum of meanings made available by a dictionary entry and its applicable encyclopedic information, the translator must choose the most probable, reasonable and relevant sense for the context in question and this possible world” (Eco, 2003)¹. The translator makes a decision for a term in a context, and the translated text becomes part of history.

Terminologists, on the other hand, must create terminological entries that are applicable in more than one situation or context, that can be comprehended by people other than the terminologist and that reflect terms used by the expert users (see Cabré, 1999). These entries are the static representations of concepts and terms in a dynamic world: Things evolve and new concepts come about; they, too, must be entered into the concept system. Terms representing these concepts are being created; they may be coined by a small circle of experts at first, entered in and standardized through a database, adopted or rejected by the community, but quickly superseded by new terms; these, too, must be added to the database (see Budin, 1996). For example, when the Office Fluent Ribbon was first created, the superordinate concept “ribbon” had not received much attention. But users and designers liked the idea, and new types of ribbons have since been created in other software programs (e.g. Windows Scenic Ribbon) and consequently made explicit in Term Studio.

If a term or definition does not stand the test against this dynamic real world—for example, if it is erroneous or simply the limited view of one person at one point in time—it invariably leads to errors. First, other, dependant entries, such as target language entries, may be affected. If an error in a target terminology entry remains unrecognized, a translator may use it and the

¹ Unauthorized translation by the author.

translation text will contain the error. The error might be propagated into other translation projects through the use of translation memory. In a very short period of time, a small inaccuracy in one entry could lead to an error in translated documents for multiple products. This would not only mean problems for product users, it would offset the return on investment (ROI) of terminology management. It is therefore of utmost importance that a complex terminology system be set up and maintained by experts who keep incorrect data to a minimum. As Cabré (1999) puts it, “terminologists should be the only professionals who actually write terminologies.”

The Terminology Process

How do terminologists arrive at sound entries that make up a correct and sustainable concept system? Data acquisition, “the process of collecting, entering, and storing data into a data processing system” (ISO 1087-2, 2000), most often starts with term extraction of source terms from one particular context. In large scale localization settings, this process is automated. The job of the terminologist, along with SMEs, is to identify valuable terms from a list of term candidates. Once a term is identified, it must be researched in other contexts. Good terminology extraction tools facilitate this process by making a variety of contexts from a variety of sources available. Furthermore, the term must be checked against existing entries in the terminology database; this, too, can be partially automated. In some cases, a more or less extensive conversation with experts about the concept and its characteristics is necessary. This research and data validation phase must also include finding superordinates and related concepts in order to avoid duplicate entries, false definitions, or other errors. Because the researcher has to identify these relations to arrive at sound entries that have certain longevity, explicating the

Barbara Inge Karsch

relations is only a matter of having the right tools functionality. Formalization and visualization of the ontology can thus be considered a byproduct of the research process that, if set up properly, can serve multiple purposes: Assist in the research of target terminology, serve as training material, facilitate search mechanisms, etc. Figure 6 shows a simple visual representation of a concept system in a terminology database in which the only goal was to support the target research process.

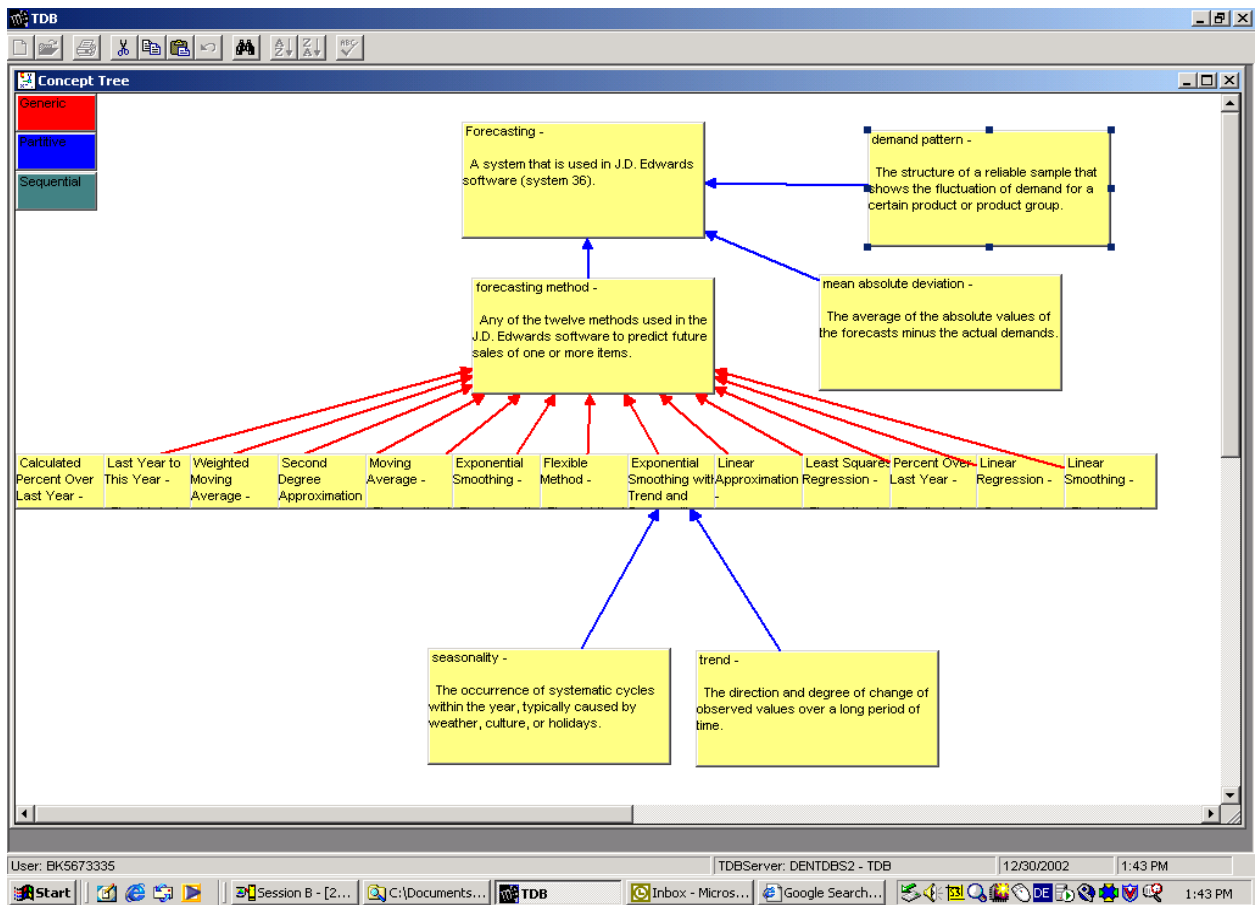


Figure 6. Visualization of a concept system for forecasting concepts in the J.D. Edwards terminology database

While this process is taking place and various data is added to the source language entry, the target term research begins. However, the target terminology process must not be concluded,

until the expert community has signed off on the definition and the conceptual and source terminology entry are correct and complete. If these steps are inverted, change management can become complicated. This process can take minutes, but it can also stretch out over months. At any given point in time, a localizer must be able to retrieve what is known and available in the system to make use of the data. Data reliability is generally indicated through the available metadata and approval statuses.

Terminologists in the Localization Industry

In their prediction for 2008, Common Sense Advisory, Inc., a research and consulting firm, expected an increase in the number of full-time terminologists in the industry beyond companies who have had terminologists on staff for many years (De Palma, 2007). In December 2008, De Palma (2008) admitted that their prediction was off. But most companies that have a high localization volume, deal with technically difficult, highly specialized terminology, or outsource to a multitude of vendors with different backgrounds have terminologists on staff. The following major computer companies were examined for this paper: SAP, IBM, Oracle, Microsoft and J.D. Edwards.

SAP has been investing in centralized terminology management the longest of the five companies. Since the early 90s, they have had a core team in charge of building up *SAPterm*, which is one of the largest terminology databases in the industry. According to coordinator for terminology issues, Mark Childress (2008), all writers and translators are responsible for terminology work as part of their tasks. Most of the product teams use the SAP board's recommendation to invest one half-day per person per week in term work, on average. By Childress's calculation, that results in 38 full-time terminologist-equivalents for the two source

Barbara Inge Karsch

languages, English and German. The vast majority of target terminology work is outsourced to translation vendors and reviewed by an in-house staff. His team of five, plus one student intern, is providing training, database maintenance, standards and processes.

The IBM structure is slightly more discreet. The team has one person to oversee the global terminology strategy, Kara Warburton, who is joined by one terminologist in charge of the source language (English), one terminologist focusing on English glossary development and management (definitions, etc.), one terminologist in charge of the English term extraction process and one part-time helper for English terminology. Furthermore, the team has one three-quarter-time project manager of terminology in the target languages and two terminology administrators in charge of importing bilingual files. There are also part-time terminologists for each of approximately 30 target languages who are focused on terminology work for a portion of their time and serve as translation project managers and translators the rest of the time (Warburton, 2008).

No staff at Oracle is solely dedicated to terminology work according to Helle Katic, Senior Process Manager on the *Oracle WPTG Translation Support Team* (Katic, 2008). Two employees have the discretionary title of terminologist, and some of their tasks are terminology-related. Furthermore, some of *Oracle's* language specialists have the role of product specialists. In that capacity, they do source terminology work on the products within their area. Likewise, language specialists do a varying degree of terminology work pertaining to their specific language. The network of language specialists consists of one or more language specialists located in the countries where translated products are marketed.

When this paper was originally delivered at Philosophy and Informatics in Fall 2008, Microsoft had five terminologists for the source language English, two full-time terminologists for Japanese, two part-time terminologists for German, and one full-time terminologist each for Brazilian-Portuguese, French, Italian, Korean, Spanish, Simplified and Traditional Chinese. Furthermore, there were dedicated vendors for 28 other languages and a team of three linguistic engineers. In spring of 2009, the terminology model was turned into a decentralized system for target languages. That means that the five source terminologists are still full-time employees, but that target terminologists are no longer available in a central team. Any product team must now hire external terminologists, who are then charged with researching and documenting target terminology.

Almost all of the localization needs of J.D. Edwards were met through an in-house team, which was established in the early 90s and existed until 2004, when the company was acquired by PeopleSoft. In 1998, a group of seven full-time terminologists grew out of the localization teams. They worked part-time on English terminology and part-time, and in cooperation with their translators, on terminology for the following tier-1 languages: Brazilian-Portuguese, French, German, Italian, Japanese, Simplified Chinese, and Spanish. All in-house translators for the tier-2 languages served as part-time terminologists documenting terms in the remaining 14 languages (see Karsch, 2002). This model was successful because knowledge sharing was possible, no complex handoff structures with vendors slowed down the process, and automated terminology workflow facilitated the complex data acquisition and entry negotiation process between stakeholders on site and in the subsidiaries abroad.

Barbara Inge Karsch

Profile of a Terminologist in the Localization Industry

The details of the processes and the complexity of the TMSs used in these companies may vary. One may be based on a quantitative approach; another has a high focus on customer satisfaction. One may have more access to SMEs; another less or not at all. One may be able to halt the localization process if no terms were documented; another may be completely at the mercy of the product teams. And yet, any of these environments require seemingly contradictory skills, characteristics and talents that translate into the rather complex profile of a terminologist. In the next section, each paragraph juxtaposes two such qualities that the author has found to be critical in her work as a terminologist for English and German first at J.D. Edwards and since 2004 at Microsoft.

Subject-Matter Expert vs. Generalist

Terminology management requires both subject-matter expertise and terminology management know-how. Many different practitioners have been involved in the field of terminology management. Engel and Picht (1999) distinguish two groups: On the one hand, there are the SMEs who engage in organizing the concepts and terms in their field. On the other hand and much more frequently, the localization industry attracts linguists who are not specialized in a subject matter, but are generalists and terminology management experts. They must be able to work with SMEs and get into a subject-matter area at the spur of the moment. For example, as part of the research for *Axapta 5.0*, a business software product from Microsoft, the terminologist dealt with terms such as *MICR line*, a line of characters that is encoded on bank checks with a special type of ink that can be magnetized and then translated into characters. Within the same week in the framework of the developer product, *Visual Studio 2008*, she dealt

with *thunk* (“A small section of code that performs a translation or conversion during a call or indirection...”). If terminologists work in teams, they can complement each other’s strengths. On the German localization team at J.D. Edwards, a certified accountant, a native Swiss with solid linguistic skills, was in charge of research and documentation of the financial terminology for German. His subject-matter expertise was rounded out by the terminology-management skills of the full-time terminologist.

Researcher vs. Negotiator

A terminologist must be able to shift quickly from focused research work to agile leadership of negotiations. Terminologists, like translators, must have excellent research skills and be informed about research material, resources, and experts; i.e. they must stay abreast of the subject-matter fields they work in. Furthermore, they must decide swiftly which resource is to be drawn on for a particular term. Research work is characterized by focus and attention to detail and at least partially carried out in a more solitary environment. Negotiations that may go along with determining the term for a particular concept, on the other hand, require a highly interactive work mode. Ideally, the terminologist leads or facilitates the negotiations between different stakeholders (see Figure 7), the outcome of which is documented in the database. This means that terminologists carry out thorough investigations first, and then, when necessary, switch into negotiation mode with their circle of experts.



Figure 7. The terminologist as the negotiator among stakeholders

‘Fingerspitzengefühl’ vs. Thick Skin

Research and negotiations not only call for the ability to work in different modes, but they also require sensitivity and thick skin. Term research or creation requires what is referred to as *Fingerspitzengefühl* in German, i.e., the accurate evaluation of resources, the meticulous creation of neologisms, and the careful guidance of the experts involved. If accuracy, precision, and sensitivity drive the research and design aspect of terminologists’ daily tasks, they must be thick-skinned in others. Language as a communication vehicle that is close to every human being is also one that is criticized quickly. Because terminologists generally have to make dozens of decisions each day, some of their choices are inadequate. Criticism is not always geared towards finding better solutions, but critical remarks could be instinctual reactions from humans, who are by default all experts in language. One important counterbalance to the sensitive creativity of a terminologist is the filter for inadvertent or deliberate disapproval.

Self-critical vs. Assertive Behavior

Critiques from others can be extremely helpful, but only work if the terminologist is open to it. A healthy dose of self-evaluation of one's terminological decisions is therefore advantageous. Reflection of one's own work and the readiness to admit that one is not infallible seem juxtaposed to the assertiveness that it takes to negotiate with others who may be experts in their field, but not aware of term formation rules, of budget constraints in terminology change management, or the requirement of a term to work for different audiences. For example, at J.D. Edwards the German term for fiscal year was changed twice from *Geschäftsjahr* to *Finanzjahr* and back to *Geschäftsjahr* again, each time costing more money. When a terminologist was appointed, she determined that the difference is merely a personal preference by SMEs and insisted that the term not be reverted back to *Finanzjahr*, although a new financial expert preferred it.

Native-language vs. Foreign-language Expertise

Terminology management requires native language experts and a tremendous understanding of the foreign language nuances. Existing terminology must be thoroughly researched, existing synonyms must be carefully weighed, and new terms must be well motivated and created under consideration of target-language term formation rules. The level of knowledge it takes to work well with one's native language is often underestimated, and yet, even the best linguists can only find good terms in their language if they are also well-versed in the nuances of the source language.

Barbara Inge Karsch

Linguistic and Technology Expertise

Linguistics and technology are two fields that are not necessarily complementary. More than a basic understanding in both is needed for a career as terminologist in the localization industry. Twenty years ago, applied linguistics or linguistics attracted students with an interest in literature and language. Database technology did not necessarily fit with that. Today, most terminologists have a good understanding of both, either because as pioneers in this field they have participated in the design, implementation, or training of terminology databases, or because the tools are still so complex that a terminologist must understand more than just how to enter data.

Organizer with a Tolerance for Chaos

By design, a terminologist is an organizer, and by definition where things need to be organized there is chaos. Even in the best-managed software projects language is not the number one priority. In a worst case scenario, software is handed off to localization and feature names or other terms are not final. The reason for this is that most software vendors are expected to ship source- and target-language versions at the same time. A terminologist can be caught in a project with not-yet-finalized terms, poorly defined concepts, conflicting definitions, or unclear concept relations. The systematic manner that a terminologist brings to the task can make a difference to the project, if the terminologist can live with temporary chaos.

Moving Quickly and Lasting Long

Sometimes a terminologist must act extremely swiftly, and at other times they must have the long-term in mind. Eliminating a mistake quickly, clarifying an ambiguous definition for a concept promptly, or taking the path of least resistance to get a term candidate approved can

make the difference for localization projects. Terminologists work with dozens of terms on a daily basis and must move quickly with decisions, find answers rapidly, and can't waste time with faulty research approaches. And yet, on occasion they need to invest time for the future. A marketing team, for example, may need to have an explanation of term formation rules in order to approve a suggested term, or a writer may need an introduction to concept relations and hierarchies to rewrite a definition. These are investments in the future of working with teams and individuals that prepare the way for quicker decisions for the next project. Sometimes, they also need to outlast an engineer to get certain functionality integrated in the TMS.

Quantity vs. Quality

During a translator's training, the emphasis is, with good reason, on quality work. In professional practice, quantity (which could be equated with on-time delivery of a job) is critical. As in many fields, the key is to deliver the highest possible quality on time. Some individuals are better in producing database entries that meet a set quality bar (e.g. the entry contains all mandatory data and is correct to the best of the terminologist's knowledge). But they may struggle with the number of entries that need to be dealt with. Others can handle the volume, but their entries may lack data, contain confusing information, or contradict other entries. In reality, though, a terminologist has to be able to spend enough time on an entry to set it up correctly, provide all necessary information, and do this consistently enough to actually respond to the needs of the database users.

Big Picture vs. Focus on Detail

When the research work necessary to set up a terminological entry is very detail-oriented, explaining what terminologists do or enlisting management support requires big-picture thinking.

Barbara Inge Karsch

Because there may only be a few terminologists on the staff of a company, they have to be able to cover both. Understanding and documenting an obscure concept requires painstaking research; often, terminologists start out with their intuition which guides them into a certain direction; until hunches are proven wrong they follow the path to find enough confirmation to conclude the research. Often, a roadblock points in another direction and more evidence must be collected, until a concept or term can be documented with enough confidence. Because of the niche-nature of the profession, business reviews happen fairly frequently, and anyone on a small team may be called upon to explain to management why terminology work is important and why funding must continue. In these situations, a business manager does not want to hear about the successes of researching the murky concept as explained earlier. Many times, they want production numbers and cost savings. Unless a terminologist can get from the detail-level research world into the big-picture world of management, communication will not happen and funding may be cut.

Theoretical Foundation vs. Pragmatic Action

Finally, there is the necessity of a good theoretical basis and the need for quick pragmatic decision-making. Experienced terminologists combine both—they can solve new business problems to serve, for instance, the majority of users quickly. But they also base these decisions on the theory that must underlie most of their actions. Pragmatic behavior without the understanding of terminology theory can lead to short-lived database entries or ill-designed TMSs. A theoretical approach without common sense will have most terminologists drowning in work. There are still many aspects of terminology management that have not been dealt with in

literature, by ISO standards or at universities. But a good understanding of the theoretical underpinnings of terminology management is a prerequisite to a successful career in the field.

Summary

Complex localization scenarios rely heavily on terminology management. If ROI is to be maximized, systematic terminology management must be based on sound concept analysis and can have as a byproduct explicit and formalized ontologies. The prerequisites for success of such a terminology and potentially ontology system are not easily satisfied—processes, tools, and people must come together in a highly complex interplay.

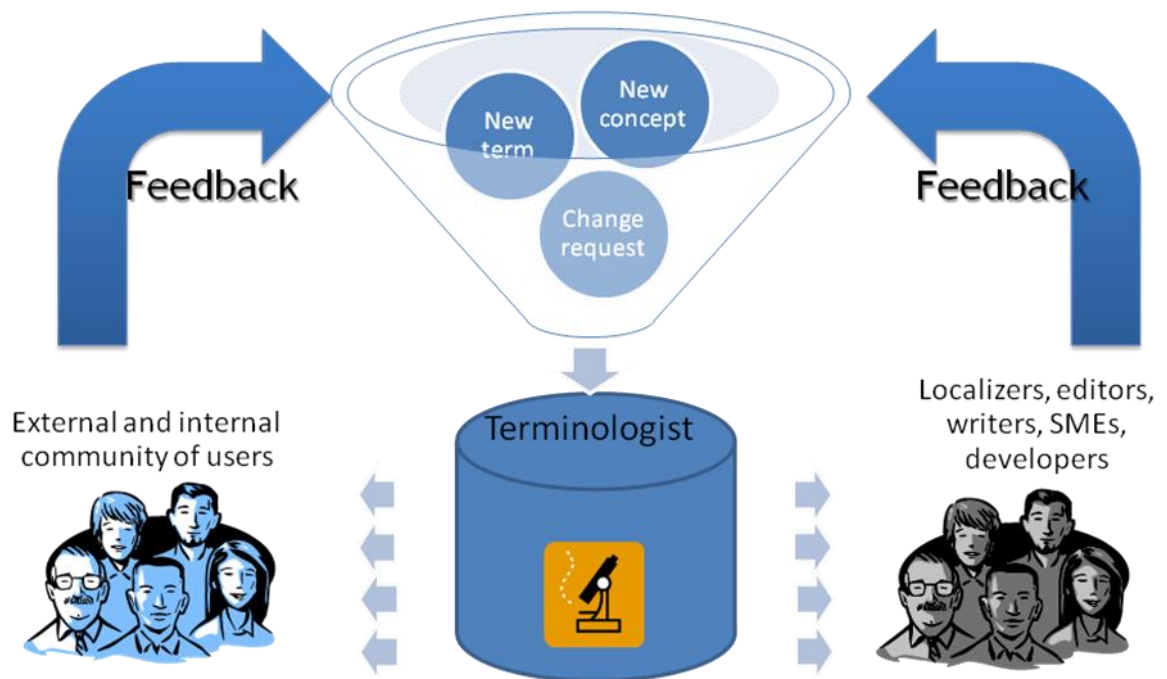


Figure 8. Simple version of a complex system in which many people contribute and many people retrieve data

Figure 8 shows such a workflow, in which ideally many contributors, such as content publishers, SMEs, or user communities funnel their knowledge into a good TMS as soon as

Barbara Inge Karsch

concepts and terms have reached a certain level of stability. Online or instructor-lead training can help enable this process, as can user-friendly input interfaces. Once entries have been created or amended, highly skilled terminology experts, who act as gatekeepers, retrieve this input. Besides their terminology processing expertise, they rely on automated processes to check validity of the information and then approve or reject entries. Once released, entries should be made available to as many users as possible. Not only does this warrant a high degree of standardized usage, it also allows for the user community to act as quality assurance. Because humans will always make mistakes even with the assistance of the most sophisticated automated processes, feedback mechanisms are vital in increasing the value of the terminology asset.

The skills it takes to set up correct entries, to keep the conceptual system sustainable, and to maintain a terminology database are manifold. Not all of the described scenarios will come up every day. Simple scenarios can be dealt with using simple solutions—a spreadsheet and one owner who collects and maintains data and distributes it back to the users. But in environments with many products and many domains, multiple languages, virtual teams, and complex tool usage, terminologists have to live up to the challenges described in this paper on a daily basis. The better equipped they are through their educational background and experience, the more successful they and the TMS will be.

Acknowledgements

The author gratefully acknowledges the contributions of Mark Childress, Helle Katic and Kara Warburton. Furthermore, she would like to thank Robin Lombard for her feedback and continued support.

References

- Beninatto, R. S., & DePalma, D. A. (2008). *Ranking of Top 25 Translation Companies*. Retrieved Nov 25, 2008, from http://www.common senseadvisory.com/members/res_cgi.php/080528_QT_2008_top_25_1_sps.php
- Budin, G. (1996). *Wissensorganisation und Terminologie: die Komplexität und Dynamik wissenschaftlicher Informations- und Kommunikationsprozesse* (Vol. 28). Tübingen: Gunter Narr Verlag.
- Cabré, M. T. (1999). Terminology: theory, methods and applications (J. A. DeCesaris, Trans.). J. C. Sager (Ed.). In: H.B. Sonneveld, & S.E. Wright (Eds.) *Terminology and Lexicography Research and Practice* (Vol. 1). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Childress, M. (2008). *SAP Terminologists*. Unpublished e-mail.
- De Palma, D. A. (2008). Predictions for 2008: How Right Were We? *Global Watchtower*TM. Retrieved Dec 23, 2008, from <http://www.globalwatchtower.com/2008/12/23/post-mortem-predict-2008/>
- De Palma, D. A., & Beninatto, R. S. (2007). Predictions for 2008: Business and Website Globalization, Technology, and Business Models. *Global Watchtower*TM. Retrieved Dec 23, 2008, from <http://www.globalwatchtower.com/2007/12/20/predictions-2008/>
- Eco, U. (2003). *Quasi dasselbe mit anderen Worten. Über das Übersetzen* (B. Kroeber, Trans.). Munich Vienna: Carl Hanser Verlag.

Barbara Inge Karsch

Engel, G., & Picht, H. (1999). *Der Terminologe - Beruf oder Funktion?* In L. Hoffmann, H.

Kalverkämper & H. E. Wiegand (Eds.), *Fachsprachen - Languages for Special Purposes*

(Vol. 2, pp. 2237-2244). Berlin, New York: Walter de Gruyter.

International Organization for Standardization (2000). International Standard ISO 704,

Terminology work – principles and methods (Vol. ISO 704:2000 (E)). Geneva.

International Organization for Standardization (2000). International Standard ISO 1087-1,

Terminology work - Vocabulary - Part 1: Theory and application (Vol. ISO 1087-1:2000

(E/F)). Geneva.

International Organization for Standardization (2000). International Standard ISO 1087-2,

Terminology work - Vocabulary - Part 2: Computer applications (Vol. ISO 1087-2:2000

(E/F)). Geneva.

Karsch, B. I. (2002). Integrierte Terminologieverwaltung - Die Einführung einer

Terminologiedatenbank beim Softwarehersteller J.D. Edwards. In F. Mayer, K.-D. Schmitz

& J. Zeumer (Eds.), *eTerminologie - Professionelle Terminologearbeit im Zeitalter des*

Internet (pp. 165-178). Cologne: Deutscher Terminologie Tag e.V.

Katic, H. (2008). *Oracle Terminologists*. Unpublished e-mail.

NeOn consortium (2008). NeOn Glossary of Activities. Retrieved Dec 12, 2008, from

<http://www.neon-project.org/web-content/images/Publications/neonglossaryofactivities.pdf>

Nistrup Madsen, B., & Erdman Thomsen, H. (2008). *Terminological Principles Used for*

Ontologies. Paper presented at the 8th International Conference on Terminology and

Knowledge Engineering.

- Reiß, K. (1986). *Möglichkeiten und Grenzen der Übersetzungskritik. Kategorien und Kriterien für eine sachgerechte Beurteilung von Übersetzungen* (3 ed. Vol. 12). Munich: Max Hueber Verlag.
- Schmitz, K.-D. (2007). Indeterminacy in software localization. In B. E. Antia (Ed.), *Indeterminacy in Terminology and LSP: Studies in honour of Heribert Picht* (Vol. 8, pp. 49-58). Amsterdam/Philadelphia: John Benjamins Publishing Company.
- SDL (2009). Terminology: An End-to-End Perspective. Retrieved Jan 30, 2009, from http://www.lspzone.com/en/Images/rp_DL%20Terminology_%20An%20End%20to%20End%20Perspective_web_tcm11-14026.pdf
- Smith, B., Kusnierczyk, W., Schober, D., & Ceusters, W. (2006). Towards a Reference Terminology for Ontology Research and Development in the Biomedical Domain. Retrieved Jan 31, 2009, from http://ontology.buffalo.edu/bfo/Terminology_for_Ontologies.pdf
- Warburton, K. (2008). *IBM Terminologists*. Unpublished e-mail.

Globalization, Reterritorialization, and Marketing

Sudhir H. Kale & Natalina Zlatevska

Bond University

Abstract

Accelerated globalization has dramatically altered the ways in which people consume, work, gather information, play and define their identity. Most extant discourse on globalization, particularly in the business discipline, ignores the impact of globalization on the identity of those affected. One of the key characteristics of globalization is deterritorialization; the severance of social, political, and cultural practices from their native places. Deterritorialization potentially destabilizes people's identity. In response, individuals will undertake activities and behaviors which help them "reterritorialize" and restore their sense of identity. This phenomenon has interesting implications for researchers as well as practitioners.

Few topics invoke as much passion amongst social scientists as the issue of globalization. Various views are held as internationalization, universalization, liberalization, westernization, and deterritorialization (Scholte, 2000), the on-going discourse and debate on globalization shows no signs of abating. Deterritorialization signifies the severance of control, order, and cultural practices away from a land or place (territory) that is already established. Potentially, the “deterritorialization” perspective should be of profound interest to scholars of international marketing and consumer behavior as it directly impacts consumption processes the world over. Yet, discussion on the deterritorialization phenomenon and its impact on consumers’ mindset and behaviors has been largely absent from the marketing literature except for a few notable exceptions (see Arnould & Price, 2000; Kale, 2004).

Scholars contend that forces of globalization have brought about profound changes in the way consumers view their identity (see Castells, 1997; Howes, 1996). It has been argued that prior to the massive onslaught of globalization that started around two decades ago, most social and cultural practices were fully embedded in an identifiable physical territory. The territory or place provided its inhabitants with a sense of stability and location and acted as a key self-descriptor. One of the main consequences of this recent wave of globalization is that people are increasingly deprived of their place-bound isolation and territorial cohesion. This creates a void in cultural and topical grounding, promoting many to seek deliberate alternative measures to restore their identity. One approach to identity restoration is through selective consumption. Indeed, evidence continues to accumulate suggesting that the frequency of purchase and use of products providing the utility of reterritorialization are on the rise the world over (Arnould & Price, 2000; Kale, 2004).

Such culturally oriented consumption runs counter to the uniformity hypothesis prophesied by scholars such as Ted Levitt (1983) and, more recently, Thomas Friedman (2006).

With the convergence of technology and income, these commentators had argued, the preferences of consumers the world over will become irrevocably homogenized. The evidence available thus far does not support their position. In the few locations where convergence of technology and narrowing of income differences have recently occurred, differences in consumer preferences have not lessened. If anything, the convergence of income has contributed to a higher manifestation of value differences (de Mooij, 2000; de Mooij & Hofstede 2002).

This article integrates literature from a diversity of disciplines – psychology, economics, sociology, and marketing, to arrive at a better understanding of consumer behavior in the contemporary global environment. We shall discuss the impact of on-going globalization on individual identity and proceed to explain how the urge to reterritorialize manifests into product preferences and use. Managerial and research implications of this perspective are then discussed.

Conceptual Underpinnings

One of the key premises of modern consumer behavior is that people often buy products not for what they do, but for what they mean (Solomon, 2004). Basically, we employ consumption not only to create and sustain the self, but also to locate us in society (Elliott, 1994). Meaning associated with products and brands is very often grounded in culture and lifestyle. In the face of deterritorialization, consumers are choosing goods and services where the perceived meaning of choice and the act of consumption fosters their sense of identity. We base this article on the premise that accelerated globalization results in increased deterritorialization, which, in turn, leads to feelings of insecure identity among consumers. Deprived somewhat of their identity, people strive to restore their sense of self by engaging in identity-enhancing consumption practices. We shall now proceed to elaborate on the foundations of our premise.

Deterritorialization

Deterritorialization causes the tearing apart of previously stable social structures, relationships, settings, and cultural representations. It constitutes the disembedding of humans and cultural symbols from their place of origin or belonging. The *Globalization Website* (n.d.) defines deterritorialization as, “Expansion of interaction and relationships not tied to or dependent on particular localities; reduced attachment to place or decreased identification with neighborhood or country resulting there from.” Scholte (2000) views globalization as the spread of supraterritoriality and a reconfiguration of geography. According to him, globalization tends to obliterate the link between social space and physical territory, thus rendering physical distances and territorial borders largely irrelevant.

Deterritorialization entails detachment of social and cultural practices from specific places, thereby blurring the age-old natural relationship between culture and geographic territories. A ‘territory’ is understood as the environment of a group (e.g., pack of wolves, a tribe, or a herd of elephants). It is constituted by the patterns of interaction through which the collective secures a certain stability and location (Kale, 2004). The environment of a single person (the social environment, personal living space, and lifestyle) can also be seen as a territory in the psychological sense, from which the person acts and returns to. In times of modernity, territorialization involved a superior power (typically the state) excluding or including people within geographic boundaries and controlling transboundary access and exchange. Such territorialization, though restricting trade and hampering internationalization, provided many citizens of the state with stability and a feeling of being centered. Territorialization through the auspices of a nation state has by no means been universally successful. As one reviewer pointed out, “Unfortunately, there are territories that were established by superpowers, with artificially defined borders that ignored pre-existing ethnic and regional turfs.” In these ill-fated

circumstances, identity was often vested not in the country, but in one's tribe or ethnic community. Regardless, the move away from modernity and toward contemporary globalization has, for millions, weakened the salience of the nation state, tribe, or ethnic heritage, and in so doing, has undermined the territorialization mechanism physically and psychically.

Globalization threatens the heretofore robust and culturally sustaining connections between geographical place and cultural experience. Globalization, thanks to transnational flows, also deprives people of their isolation and territorial cohesion, thus causing many to question their identity or sense of being. The massive flow of people, capital, media, and commodities in and out of previously sovereign territories means that the long-standing relationships between culture and place are getting increasingly disconnected (Appadurai, 1996; Hannerz, 1989). Consequently, many people sense a deep void in cultural and topical grounding. The ensuing rootlessness, for many, creates a crisis of identity and an urgent, often frantic urge to reterritorialize. Attempts to reterritorialize in contemporary postmodern times are often symbolic with people wanting to hold on to whatever they feel most defines their identity.

Deterritorialization prompts changes in attitudes as well as behaviours, resulting in a culturescape that may be global and cosmopolitan in some aspects, but deeply tribal and territorial in others. For, as globalization intensifies, it manifests deterritorialization, which in turn, evokes identity preserving responses through attempts at reterritorialization. Buchanan (2004, p. 17) captures this phenomenon cogently when he writes, "...although most of us embrace the opportunities globalization affords us, we nonetheless continue to sense and long for a past none of us has actually known when the connections were local not global, when the food on our plate was the result of our own toil in the garden. This is the world, as imaginary as it obviously is, that we have been evicted from by our own success in transforming our habitat. The longing underpinning this feeling of exile manifests itself in the form of disorientation, we

can't seem to get our bearings in this brave new world without borders. Disorientation brought about by the disembedding process requires in turn a compensating process of reembedding to accommodate us to the alienatingly 'faceless' world of modernity."

Impact on Consumers

New communication technology such as telephony and the Internet, massive migration, and the dissolution of the iron curtain have drastically increased the pace of globalization, accompanied by massive deterritorialization and a resulting urge to reterritorialize. Appadurai (1996) was among the first to explore the cultural dimensions of globalization. In explaining how present-day globalization is different from earlier global movements, he observes that contemporary globalization has shrunk the distance between elites, shifted the dynamics of exchange between producers and consumers, broken many links between labor and family life, and most importantly, obscured the lines between temporary locales and imaginary national attachments.

As mentioned earlier, deterritorialization applies not only to transnational corporations, money markets, and technology, but also to ethnic groups, sectarian movements, political formations, and individual identities. Deterritorialization has resulted in enormous fragmentation in the constitution of people's demographics, tastes, preferences, and self-concept in any given geographic unit (city block, town, state or nation). Firat (1997, p.78) observes, "Fragmentation is reflected in the simultaneous presence of different and essentially incompatible patterns and modes of life represented by a variety of products, lifestyles, and experiences that do not fit with each other, instead representing different cultural identities and histories."

Contemporary consumption is increasingly undertaken to rediscover, preserve, or even create one's identity. Several scholars have discussed the relationship between various facets of globalization and identity. Baumeister (1996) explains how detraditionalization increases choice

and, in so doing, problematizes identity formation. McDonald (1999) has argued that many young people feel marginalized by globalization as their identities can no longer be constructed within the imagery and culture created by producers and employers. Touraine (1997) relates fragmentation and loss of identity to demodernization, and, likewise, Kayatekin and Riccio (1998) relate the loss of self to globalization.

According to Erikson (1959), identity basically implies perception of one's continuity and coherence in time and it comprises of genetic and cultural components. Identity crisis is one of the most severe conflicts that humans encounter. In underscoring the role of identity, Erikson remarks, "In the social jungle of human existence, there is no feeling of being alive without a sense of identity" (quoted in Gergen, 1991, p. 38).

It is not just genetic evolution that finds expression in identity. The evolution of memes or transferable ideas is also incorporated in a person's sense of self. These ideas involve cultural values and practices, most notably religion and language. Until recently, culture and religion were largely territory-specific. The forces of technology, media, and migration have served to dislocate these ideas from their geographic origin and often times defiled their core, thus creating an identity crisis in many of the world's inhabitants. Castells (1997, p.3) explains,

In a world of global flows of wealth, power, and images, the search for identity, collective or individual, ascribed or constructed, becomes the fundamental source of social meaning. This is not a new trend since identity, and particularly religious and ethnic identity, have been at the roots of meaning since the dawn of human society. Yet identity is becoming the main, and sometimes the only source of meaning in a historical period characterized by widespread destructuring of organizations, delegitimation of institutions, fading away of major social movements, and ephemeral cultural expressions. People increasingly organize their meaning not around what they do but on the basis of who they are.

Arnould and Price (2003) have observed that under globalization, key reference points for identity like community, nation, and people become fluid and contentious, thus rendering global culture "contextless" (Smith, 1990). Similarly, Featherstone (1991) has argued that we are

increasingly moving to a global society in which adoption of previously fixed and territory bound lifestyles for specific groups has been surpassed.

The beseeching quest for identity in the throes of deterritorialization finds significance in consumer behavior. Increasingly, the act of consumption becomes an important vehicle with which to shape one's self-image and define one's identity. Arnould and Price (2003) explain that contemporary consumers, in the midst of multiple identities that globalization affords, and lacking authority and continuity formerly provided by tradition, actively hunger for a sense of continuity and integration. Consumers, through creative consumption, seek to create a foundation whereby their authentic selves and connection to community could be reclaimed. In doing so, individuals strive to combat the tyranny of deterritorialization by engaging in reterritorialization to guard their identity. Appadurai (1996, p. 4) cites media consumption patterns to illustrate reterritorialization by diasporas, "As Turkish guest workers in Germany watch Turkish films in their German flats, as Koreans in Philadelphia watch the 1988 Olympics in Seoul through satellite feeds from Korea, and as Pakistani cabdrivers in Chicago listen to cassettes of sermons recorded in mosques in Pakistan or Iran, we see moving images meet deterritorialized viewers."

Reterritorialization motives are not simply confined to the consumption and media habits of diasporas. As Arnould and Price (2003) and Kale (2004) have discussed, consumers *everywhere* are aspiring to reterritorialize; reclaiming markers of identity wherever they could be found – in tribalism, nationalism, ancestry or religion. Consumption as reterritorialization helps explain the resurgent demand for local products among residents from the former East Germany. Dornberg (1995/1996) reports that East German brands for wine, spirits, beer, cosmetics and detergents have all made a comeback thanks to renewed patronage from consumers in the East. Referring to the unification of Germany and the subsequent nostalgia (called 'Ostalgie' in

Germany), Alexander Mackat, an ad agency executive catering to Eastern consumers explains, “We were told our biography is horrible and has no meaning. We had to cocoon ourselves for self protection, and this was the beginning of ostalgie” (Fitzgerald, 2003). This nostalgic feeling is not confined to those who actually grew up in the former German Democratic Republic. Amazingly, it is also shared by their children, most of whom have only a vague memory, if any recollection at all, of the GDR.

Religion and Global Business

Scholars such as Levitt (1983) considered globalization a mechanism of standardization that would transcend regional culture, religious practices, and tradition. This scenario has been described in several clichéd terms such as McWorld, McDonaldization, Disnification, and Coca-colonization. A corollary to the homogenizing influences of globalization is secularization. Sociologists in the 1970s, then looking ahead to the year 2000, regarded secularization as the most relevant religious scenario for the end of the century (Introvigne, 2004). Statistically, a large body of empirical research has now conclusively established that religion is more prevalent in the 2000s than it was in 1970s (Stark & Finke, 2000). The number of persons calling themselves “religious” has increased in almost all countries in the world (Introvigne, 2004).

The renewed and reinvigorated appeal of religion over the past three decades could be attributed to deterritorialization (Kale, 2004). For millennia, both religion and nationhood had been deeply rooted in geography. The loss of national ethos, a direct consequence of globalization, may be the reason why billions of people are now looking at religion as an alternative marker of identity.

Of the over 150 major conflicts that occurred from 1990-2000, only two could be considered truly international. Over thirty such conflicts precipitated major wars which claimed the lives of over 40 million people and dislocated 45 million more. Religious and ethnic

differences fuelled an overwhelming majority of these conflicts (Kale, 2004). In tracing the roots of several recent ethnic and religious conflicts to deterritorialization, Appadurai (1998, p. 919) concludes, “To put it in a sanitized manner, the most horrible forms of ethnocidal violence are mechanisms for producing persons out of what are otherwise diffuse, large-scale labels that have effects but no locations.” In other words, the macabre ethnic killings, to many, serve as a means to reterritorialize. Appadurai (1998, p. 920) goes on to explain, “... as large populations occupy central social spaces and as primary cultural features (clothing, speech style, residential patterns) are recognized to be poor indicators of ethnicity, there tends to be a steady growth in the search for ‘inner’ or ‘concealed’ signs of a person’s ‘real’ identity. The maiming and mutilation of ethnicized bodies is a desperate attempt to restore the validity of somatic markers...” It is often in terms of the “other” that individual identity can be recognized, and causing harm to this “other” can sometimes be viewed as a pathway to reclaiming one’s identity.

As ethnic and religious conflicts multiply, global firms are likely to face unprecedented instability in the various markets they operate. Such hostility is normally imagined within the context of remote and ‘backward’ areas of the world, which only serves to undermine the seriousness of the issue. Cultural pride was at the roots when French men and women attacked a *McDonald’s* restaurant for its assumed assault on French culinary culture. Liberalization and tearing down of trade barriers continues to open more and more markets for global companies, but the rising identity crisis could render many of these markets politically unstable and physically unsafe.

The reterritorialization movement demands that global companies exercise greater sensitivity to ethnic preferences in product design--particularly in relation to food, clothing, and recreation. Not doing so could produce a hostile and often violent backlash. Catering to the needs of deterritorialized customers is a tricky proposition. If identity-affirming consumer

innovations (or revisitations) in dress, food, or lifestyle are re-packaged and mass-produced, these acts may be perceived as further evidence of cooptation of ethnic symbols for profiteering. On the other hand, not doing so may be attributed to cultural insensitivity and to the imperialistic orientation of global firms.

Global firms will increasingly have to contend with not only reterritorializing consumers but a reterritorializing workforce as well. The phenomenal increase in books and seminars on spirituality in the workplace is but one manifestation of this. Arnould and Price (2003), on the basis of qualitative interviews with people from several countries, have demonstrated increased incidence of “authenticating acts” and “authoritative performances” to restore a sense of community, tradition, and self. Religion and spirituality can encompass both authenticating acts as well as authoritative performances. Today, many people are demanding that such acts be practiced in public places, including the workplace.

Authenticating acts denote self-referential behaviors that are deemed to reveal or produce the “true” self. These acts are conducted in search of true identity and involve transcending the superficial roles played by individuals (e.g. while at work). Authoritative performances tend to be collective in nature and often involve cultural displays such as festivals and rituals, designed to create a communally shared experience. Authoritative performances represent the quintessential search for unity between self and society (Cohen, 1988). Both authenticating acts and authoritative performances serve to create a sense of “being centered” in today’s deterritorialized society. As spiritual and religious practices typically involve authenticating acts as well as authoritative performances, employees are increasingly expressing the need for their employers to accommodate such activities (Mitroff & Denton, 1999). The challenge for global companies, therefore, is to create a unifying global culture, while simultaneously catering to the diverse identity needs of the workforce.

When viewed against the backdrop of deterritorialization, recent practices of outsourcing customer contact centers to places such as India seem particularly perverse. Call center employees are often asked to assume a persona totally foreign to them, an alien name, an artificial accent, a fabricated life story. Economic pressure often causes many employees to acquiesce to such demands. But Erikson (1959) and Castells (1997) caution us that continued and deliberate (even if forced) negation of one's true identity produces pathological consequences and dysfunctional behaviors. Unlike their employers living in gated suburban communities in the same city or thousands of miles away, call center agents – between inhumanely scheduled shifts – have to return to their run-down *chawls*, discard their Western persona, and make sense of their local territory. Not unlike uniforms of the industrial era, present day call center employees don at work a new guise prescribed by their employers, playing the part of a local for callers stationed thousands of miles away. Unfortunately, while a uniform touches but the outermost reaches of one's being, the masked identity of contact center agents cuts to the very core. Their job is undoubtedly an act, but unlike a play or a movie, audiences – the callers – are unaware of it being an act. This double deception, deception of the employee's own identity, and deception of the callers, will sooner or later, spark a schizoid personality and pathological behaviors. When the causes of these behaviors are widely understood, global companies would most likely become targets of widespread hostility.

Implications for Scholars

Researchers in business have thus far ignored the impact of globalization and deterritorialization on people's identities. Yet, as Virilio pointed out twenty-five years ago, identity is *the* major question for this century (Virilio & Lotringer, 1983). For scholars in international management and marketing, superficially studying the phenomena of localization

vis-à-vis globalization is not enough. The motives toward the localization movement need to be better understood. This is desirable from a social welfare as well as from a business perspective.

It is time for management and marketing scholars to ponder on the psychic costs of globalization and make their contributions to this vital area. Issues such as identity, heritage, tradition and spirituality can no longer be ignored. In this regard, it behoves us to remember Kierkegaard's (1989, p. 62) comment in his novel *The Sickness Unto Death*, "The biggest danger, that of losing oneself, can pass off in the world quietly as if it were nothing; every other loss, an arm, a leg, five dollars, a wife, etc. is bound to be noticed".

Management and marketing scholars interested in global issues need to better appreciate the role "place" has played and continues to play in the lives of people. Thanks to globalization, present day identities involve more complex types of mixing and dialectics than in the most recent past (Escobar, 2001). As Escobar (2001, p. 169) warns, "The fact that a growing number of people and groups demand the right to their own cultures, ecologies, and economies as part of our modern social world can no longer be denied, nor can these demands be easily accommodated into any universalist or neo-liberal doctrine". Not attending to issues of deterritorialization and place could thus prove to be a serious blind spot for scholars and educators.

Conclusion

Deterritorialization constitutes the process wherein the very basis of one's identity, the proverbial "ground beneath our feet," is eroded, washed away like the bank of a river accosted by floodwater (Buchanan, 2004). Globalization has endeavoured to create a blanket of conformity among consumers and individuals worldwide, thereby obfuscating age-old markers of identity. In response, consumers are engaging in acts of reterritorialization, individually and collectively. A movement has been born where individuals are exercising their right to selective consumption

by turning away from global products perceived as killers of identity. Consumers, increasingly, are embracing products, customs and traditions that signify and signal who they are and where they have come from. Although the typical Italian family, having migrated to Australia is open to a myriad of food choices from American style burgers, to Thai, to Chinese, come dinner time, the family, still prefers the familiar taste and aroma of home made pasta, a tradition that generations of this family to come, will still uphold.

References

- Appadurai, A. (1996). *Modernity at large: Cultural dimensions of globalization*. Minneapolis, MN: University of Minnesota Press.
- Appadurai, A. (1998). Dead certainty: ethnic violence in the era of globalization. *Development and Change*, 29, 905-925.
- Arnould, E. J., & Price, L. L. (2003). Authenticating acts and authoritative performances: questing for self and community. In C. Huffman, S. Ratneshwar, & D. Mick (Eds.), *Contemporary Perspectives on Consumer Motives, Goals, and Desires*. New York, NY: Routledge Press.
- Baumeister, R. (1996). Identity as adaptation to social, cultural, and historical context. *Journal of Adolescence*, 19 (5), 405-17.
- Bosker-Liwerant, J. (2002). Globalization and collective identities. *Social Compass*, 49 (92), 253-271.
- Buchanan, I. (2004). Space in the Age of non-Place. *Drain Magazine*, 3. Retrieved January 15, from <http://www.drainmag.com/issue03/html>.
- Buchwalter, A. (2001). Globalization, modernity, and fundamentalism. *UC Santa Cruz Currents Online*. Retrieved January 25, 2005, from http://currents.ucsc.edu/terrorist_crisis/crisis_forum.html.
- Casanova, J. (2001). Religion, the new millennium, and globalization. *Sociology of Religion*, 62 (4), 415-41.
- Castells, M. (1997). *The power of identity*. Malden, MA: Blackwell.

- Cohen, E. (1988). Authenticity and commoditization in tourism. *Annals of Tourism Research*, 15 (3): 371-386.
- de Mooij, M. (2000). The future is predictable for international marketers: converging incomes lead to diverging consumer behavior. *International Marketing Review*, 17 (2), 103-113.
- de Mooij, M. & Hofstede, G. (2002). Convergence and divergence in consumer behavior: implications for international retailing. *Journal of Retailing*, 78, 61-69.
- Dornberg, J. (1995/1996). Five years after reunification – Easterners discover themselves. *German Life*. Retrieved December 29, 2004, from http://germanlife.com/Archives/1995/9512_01.html.
- Elinder, E. (1965). How international can European advertising be? *Journal of Marketing*, 29, 7-11.
- Elliott, R. (1994). Exploring the symbolic meaning of brands. *British journal of Management*, 5, 13-19.
- Erikson, E. (1959). Identity and lifecycle. *Psychological Issues*, 1, 1-171.
- Escobar, A. (2001). Culture sits in places: reflections on globalism and subaltern strategies of globalization. *Political Geography*, 20, 139-174.
- Featherstone, M. (1991). *Consumer culture & postmodernism*. Newbury Park, CA: Sage.
- Firat, A. F. (1997). Globalization of fragmentation – a framework for understanding contemporary global markets. *Journal of International Marketing*, 5 (2), 77-86.
- FitzGerald, N. (2003). East Germans hanker after bygone things. *International Herald Tribune*, March 12, 16.

- Gergen, K. (1991). *The saturated self: Dilemmas of identity in contemporary life*. New York, NY: Basic Books.
- Giddens, A. (1990). *The consequences of modernity*. Stanford, CA: Stanford University Press.
- Globalization Website (n.d.). Retrieved from,
<http://www.sociology.emory.edu/globalization/glossary.html#D>
- Hannerz, U. (1989). Notes on the global ecumene. *Public Lecture, 1* (2), 66-75.
- Harvey, P. (2001) Landscape and commerce: creating contexts for the exercise of power. In B. Bender & M. Winer (Eds.), *Contested Landscapes*. New York, NY: Berg
- Hofstede, G. (1991). *Cultures and organizations: Software of the mind*. New York, NY: McGraw Hill.
- Howes, D. (Ed.) (1996) *Cross-cultural consumption: Global markets, local realities*. London: Routledge.
- Husted, B. W. (2003). Globalization and cultural change in international business research. *Journal of International Management, 9* (4), 427-433.
- Introvigne, M. (2004). The future of new religions. *Futures, 36*, 979-990.
- Kale, S. H. (1995). Grouping Euroconsumers: a culture-based clustering approach. *Journal of International Marketing, 3* (3), 35-48.
- Kale, S. H. (2004). Spirituality, religion, and globalization. *Journal of Macromarketing, 24* (2), 92-107.
- Kayatekin, A. & Ruccio, D. (1998). Global fragments: subjectivity and class politics in discourses on globalization. *Economy and Society, 27* (1), 74-96.

- Kierkegaard, S. (1989 [1849]). *The sickness unto death*, trans. Hannay, A. London, UK: Penguin.
- Levitt, T. (1983). The globalization of markets. *Harvard Business Review*, 61 (3), 92-101.
- McDonald, K. (1999). *Struggles for Subjectivity*. Cambridge, UK: Cambridge University Press.
- Mitroff, I. & Denton, E. (1999). A study of spirituality in the workplace. *Sloan Management Review*, 40 (4), 83-92.
- Roostal, I. (1963). Standardization of advertising for Western Europe. *Journal of Marketing*, 27, 15-20.
- Scholte, J. A. (2000). *Globalization: A critical introduction*. London, UK: McMillan.
- Smith, A. D. (1990). Toward a global culture? In M. Featherstone (Ed.), *Global culture*. Newbury Park, CA: Sage.
- Stark, R. & Finke, R. (2000). *Acts of faith: Explaining the human side of religion*. Berkeley, CA: University of California Press.
- Touraine, A. (1997). *What is democracy?* Boulder, CO: Westview Press.
- Virilio, P. (1999). *Politics of the very worst*. New York, NY: Semiotext(e).
- Virilio, P. & Lotringer, S. (1983). *Pure war*. New York, NY: Semiotext(e).

Video Game Localisation for Fans by Fans: The Case of Romhacking

Pablo Muñoz Sánchez

University of Granada

Abstract

The aim of this paper is to show the particularities of the so-called “romhacking”, a methodology developed by amateurs to localise mainly classic video games. In the first section, the concept and origin of the term “romhacking” is presented. The second section offers an overview of the workflow followed by romhackers to localise video games. In the third section, an analysis of the differences between professional and amateur translations is given. The fourth section includes a discussion of the legal aspects of this practice. The paper concludes with a reflection on the impact of amateur translations on the video game localisation industry.

Although the majority of video games are created in Japan or North America, some are never released from their country of origin, despite the fact that they yield large profits (up to 50%) in other international markets such as Europe (Chandler, 2006). Fortunately, the situation began to change in the mid-1990s (Bernal Merino, 2006, pp. 23-24), although some video games still are not localised when released in other countries. For instance, 490,000 copies of *Chrono Trigger DS* were shipped to Japan and 220,000 to North America as of December 2008 (Square Enix, 2009, p. 19), but it was released in Europe on February 2009 only in English.

On the other hand, some gamers think that today's video games are not as good as several decades ago, and many nostalgic fans have opted to relive classic video games (Gieske, 2002, pp. 9-10). Indeed, some of the authors who have written about the history of video games, such as Burnham (2001), Poole (2000) or Sellers (2001), consider the 1990s as the golden age of these entertainment products, arguing that today's outstanding graphics do not compensate for the lack of originality.

Thanks to the advent of the Internet and the so-called "console emulators" (i.e., software that allows playing games of a specific console on a different platform, usually a PC), now gamers can easily download a file containing the binary data of an old game (which is called ROM, i.e. Read Only Memory) and enjoy the old *Pac-Man* or *Space Invaders* as if they had the original hardware and software. Although the ideal aim of emulators is to preserve the cultural heritage of video games (Esposito, 2004, p. 19), gamers now find it possible to play video games that were never shipped from the country where they were developed.

However, as many popular games were released only in Japanese or English, many gamers could not have a complete, enjoyable experience. Thus, some fans started to study the structure of the ROM data of their favourite games, developing tools everybody could use to do

fan translations. These are the origins of the so-called “romhacking”, that is, the process of modifying the ROM data of a video game to alter the game’s graphics and dialogues.

As the author of this paper has been a romhacker in the past, his aim is to give an overview of the methodology followed by romhackers to localise video games and to discuss some of the quality and legal aspects involved in these amateur translations from an academic point of view.

The Concept of Romhacking

The term “romhacking” originates from the words “ROM” (Read Only Memory) and “hacking”, as romhackers need to “hack” the ROM data of a game to alter its contents. The word “hacking” refers here to the fourth definition of hacking in the *Merriam-Webster’s 11th Collegiate Dictionary* (2009): “to write computer programs for enjoyment”. In fact, according to Esselink (2002), the skills of a typical romhacker resemble those of a localisation engineer: “A localisation engineer doesn’t need to be a programmer, a translator, or a designer. Most importantly, a localisation engineer should be someone who is not afraid of technology, who loves to experiment and troubleshoot, and who is extremely flexible” (p. 4).

Generally speaking, the term “romhacking” refers to the process of modifying the ROM data of a video game to alter the game’s graphics, dialogues, levels, gameplay or other gameplay elements. A romhacker may have two kinds of projects: editing a game to create new levels and to change characters’ attributes; or translating it from one language to another. In this paper, only translation projects are addressed, as the first type is beyond the scope of this research.

Another usual term related to romhacking is “fan translation”. However, this should not be confused with actual romhacking, because fan translation is simply a translation carried out by amateurs, no matter the product, as is the case of “fansubs” (subtitles created by amateurs for

Japanese cartoons; see Díaz Cintas & Muñoz Sánchez, 2006). As explained above, a romhacking project does not necessarily involve translation at all.

The Methodology of Romhacking

Depending on the complexity and amount of text to translate, romhacking projects may involve a team of several romhackers and translators or only one romhacker who also acts as a translator. For example, the romhacking project of the *Super Nintendo* game *Tales of Phantasia* involved a romhacker and two Japanese-to-English translators from the romhacking group *DeJap* because the script was very long, while the English-to-Spanish translation and romhacking of the *Super Nintendo* game *Secret of Mana* was carried out by only one person, as the script was not that lengthy. Moreover, the text display in each game is unique to that game, so the efforts devoted to romhacking may vary from one project to another.

The following paragraphs offer an overview of a typical romhacking project, although one must note that not all games can be translated following this methodology, as sometimes the text data in a ROM is compressed with different algorithms. For more information on the process, the comprehensive guide of Muñoz Sánchez (2007) explained in detail nearly all the different techniques to extract and insert text in a ROM. The website *Romhacking.net* (2009) also provides a good deal of romhacking documents and tools (the software cited in this paper can be found on said website).

Text decryption. The text contained in the ROM data is usually encrypted, meaning that it cannot be read with software like Notepad or Microsoft Word. Therefore, a romhacker must use software such as *SearchR X* or *Search Relative* (Figure 1) to decrypt the text and create a “table”, that is, a file containing the equivalence codes of the decrypted characters in a ROM so other software designed for romhacking can decrypt and thus show the text accordingly. For this task, a romhacker needs to search for a word or phrase appearing in the game, note down the

equivalence codes of the decrypted text (which are presented in hexadecimal, that is, a numeral system with a base of 16 that uses the symbols 0-9 to represent values zero to nine, and A through F to represent values ten to fifteen).

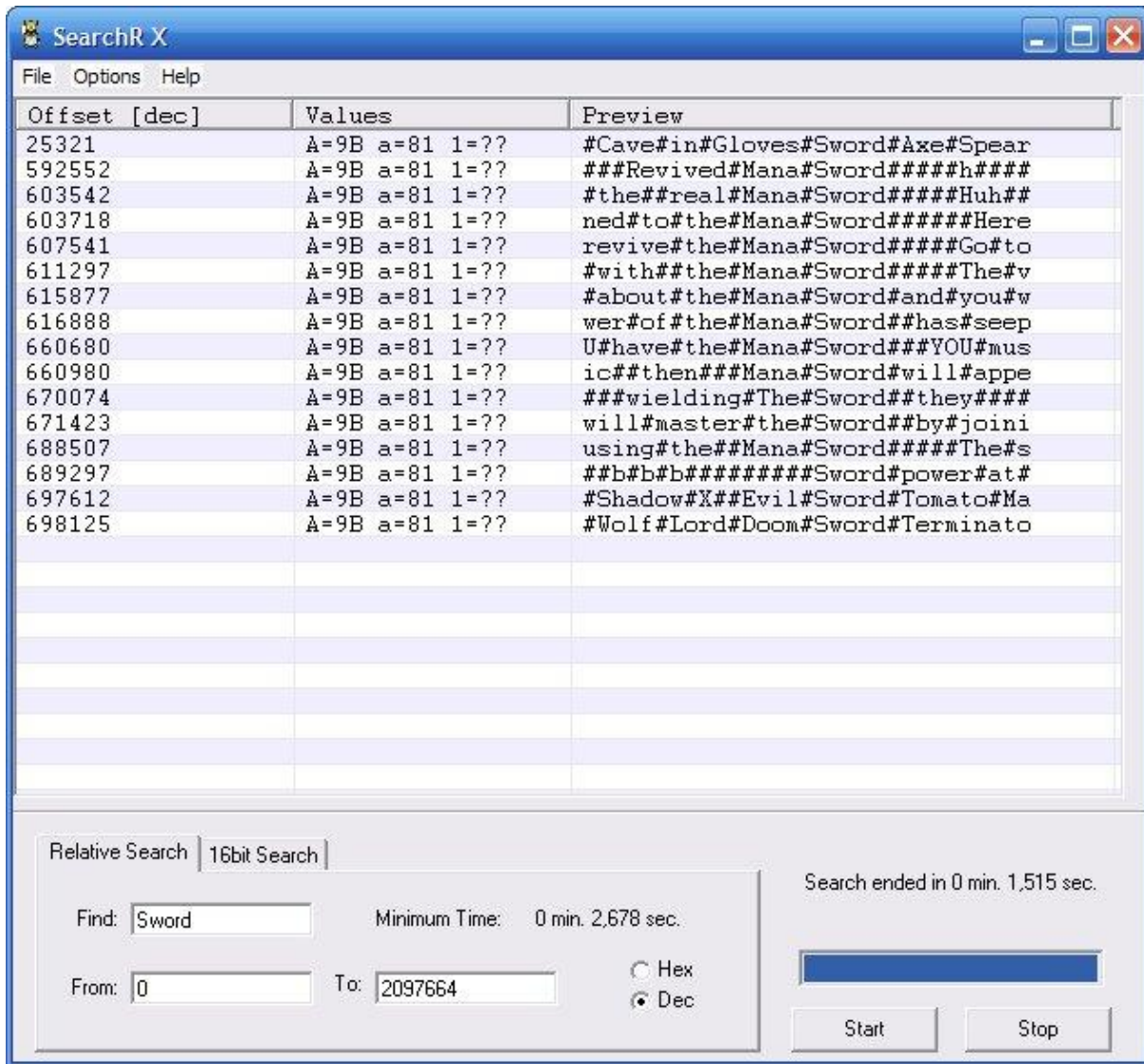


Figure 1. Searching for the word “Sword” in the *Secret of Mana* ROM using *SearchrR X*.

Once the text has been decrypted, the creation of a table file can be automated with software such as *TaBuLar* or *TAG* (Figure 2).

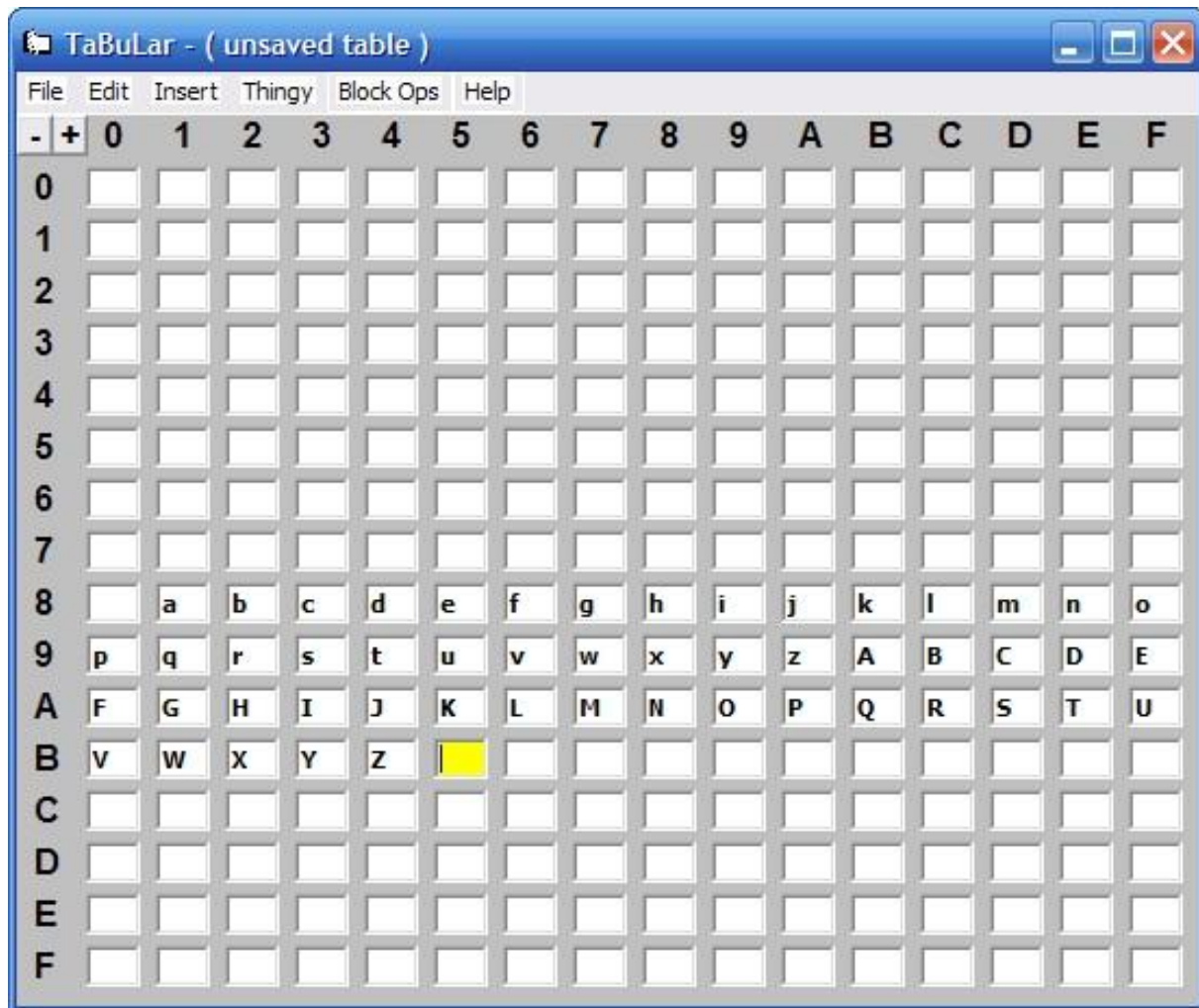


Figure 2. Automating the creation of a table file using *TaBuLar*.

Graphic edition. Normally, a non-localised ROM does not allow using accented letters, and Japanese ROMs may not even contain the Latin alphabet. Therefore, prior to the translation of a ROM, romhackers must edit its graphic data to insert all the necessary letters and symbols. In order to achieve this task, romhackers use software such as *Tile Layer Pro* or *Tile Molester* to edit unused characters in the target language (e.g., Japanese characters if the video game is in Japanese or rarely-used characters in Spanish such as the apostrophe if the video game is in English) and thus create a special character (Figure 3). This way, if a romhacker changes a

semicolon for an accented “a,” the video game will show an accented “a” instead of a semicolon when using the equivalence code for the semicolon.

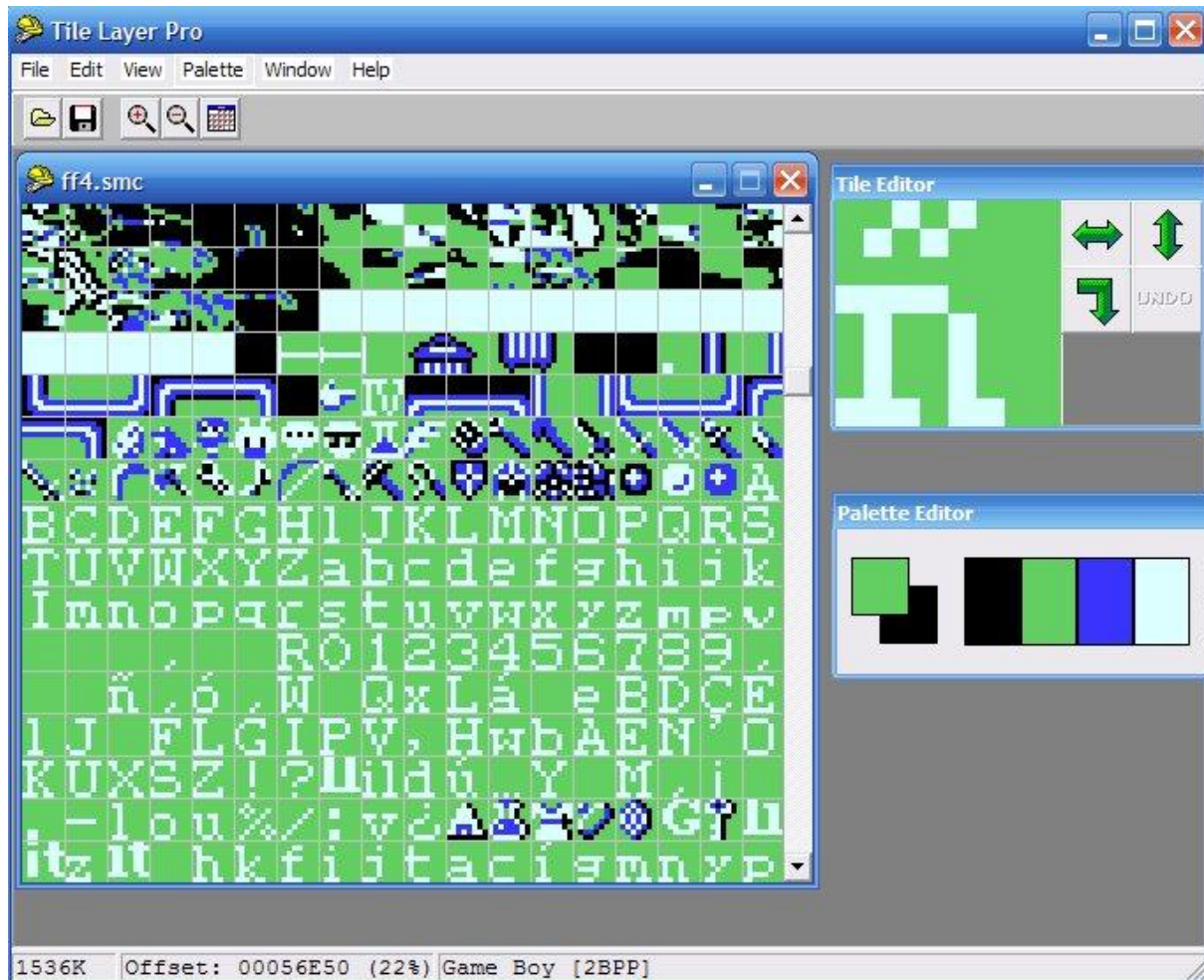


Figure 3. Font of the *Final Fantasy IV* ROM (*Super Nintendo* version) in *Tile Layer Pro*.

Nevertheless, editing graphics in a ROM is not a process similar to retouching a normal image. The type of graphic data in a ROM depends on the console the video game was designed for, so romhackers first need to set up the software according to a specific console (e.g., *Tile Layer Pro* supports the *Super Nintendo* and *Game Boy* formats, among others) and then look for the ROM font with extra care, as graphics do not usually appear in order (Figure 3). If graphics are difficult to find, trying other formats may help.

Translation. Once romhackers have decrypted the text in a ROM, created a table file with most equivalence codes, and edited the graphic data to use special characters, they can start the translation phase. Romhackers use two main approaches to carry out this process: using special software called “hex editors” together with the table file and starting the translation; or locating all the portions of the script within the ROM, exporting them to normal text files, and then importing them once the translation is finished. While the first option is faster, the second gives more control over the text and favours the translation of a large script among different translators. The main hex editors used for romhacking projects are *WindHex* and *Translhextion*; both tools allow translating on the fly as well as extracting and inserting scripts to a text file.

Before starting to translate, romhackers need to complete the table file with some missing equivalence codes, such as dots, commas, line breaks and other special tags used to display text on screen. To do so, romhackers have to load a ROM together with its table file in a hex editor designed for romhacking and search for some text appearing in the game. Around the text will be non-recognised codes, so the romhackers need to deduce what the codes mean. The best way to finish the table is to search for text appearing at the beginning of the game and then compare the differences between the text in the hex editor and the on-screen text. In order to add new equivalence codes to the table file, romhackers need to open it with a text editor, such as Notepad, and follow the format of previous entries, which is normally “hexadecimal code” = “character”. Figure 4 shows an example of this process.

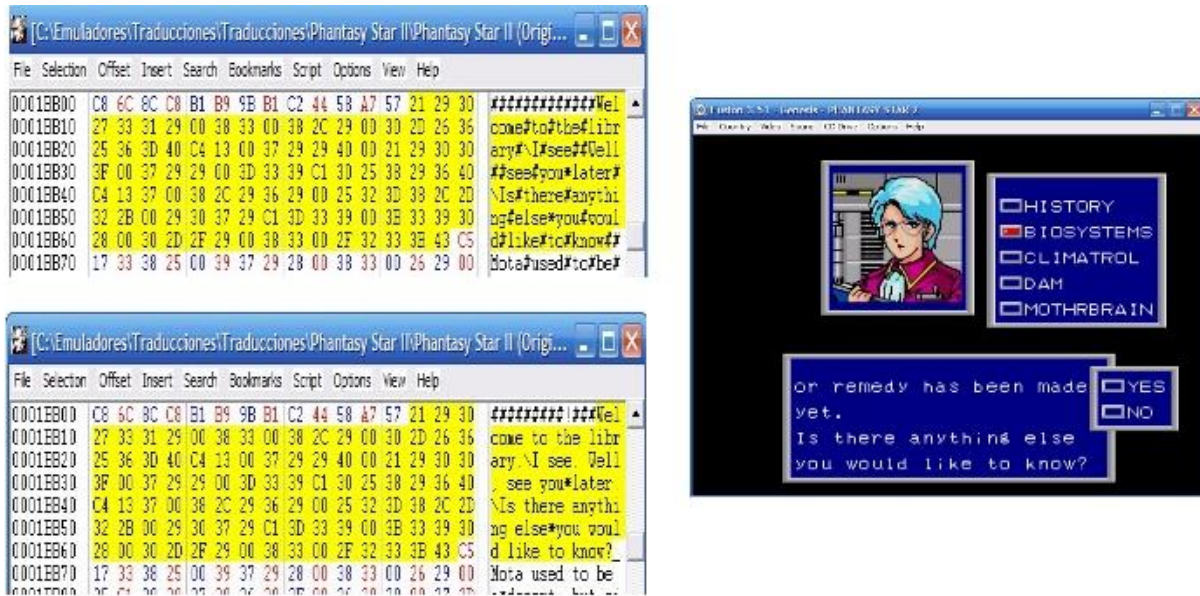


Figure 4. Completing a table file using on-screen text as reference.

When the table file is complete, a romhacker can then start translating. As using a hex editor to translate a ROM is not very comfortable, many romhackers opt to dump all the text into text files. In fact, this is what game developers do to localise their products: they prepare Excel files with all the translatable content (usually with many comments and macros to help translators) so it can be put back into the game easily once the translation is finished.

On the other hand, romhackers face a handicap when trying to achieve a good translation, as they cannot use more characters than in the original message because the text may not appear properly on screen, so concision sometimes plays a key role in the translation. However, advanced romhackers know how to locate and modify the so-called “pointers”, that is, internal codes in the ROM that specify when dialogues should be displayed. Thus, they can expand some messages and shorten others so the entire script fits into the ROM. Although software that makes this task easier is available, a romhacker must always know how to locate these pointers. Consequently, a team of expert romhackers and translators who can divide the tasks will be ideal to work on large projects and produce good quality translations.

Fortunately, as romhackers usually have programming skills, they often develop localisation kits so translators can see how the text will be displayed on screen, which can avoid many bugs after importing the translated text into the ROM, especially those related to layout and character limitations. For instance, *FF3 Editor* (Figure 5) is software that allows for editing almost everything in the *Final Fantasy III* (US version) ROM, including the dialogues and font.

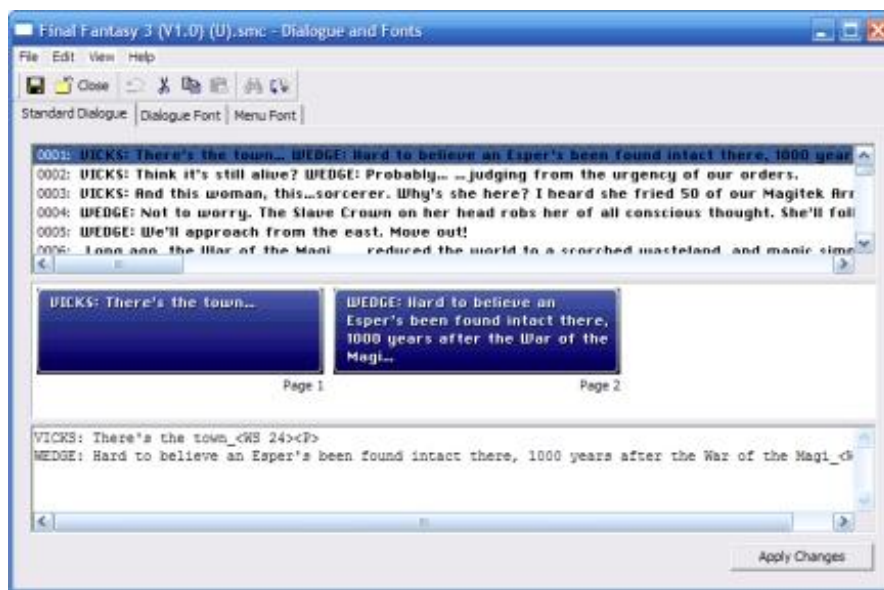


Figure 5. The dialogue editor of *FF3 Editor*.

Finally, when the translation is finished, a long testing process begins to spot and fix bugs. Once the text is considered bug-free and accurate, romhackers create a “patch”, that is, a file that can be applied to the original ROM in order to create the localised version.

Differences between Professional and Amateur Translations

Theoretically, the quality of fan translations should be below par compared to professional translations. Indeed, anything from typos to mistranslations are often found, but as in the case of fansubs (Díaz Cintas & Muñoz Sánchez, 2006, p. 45), sometimes amateurs’ efforts to be perfectionists have yielded good results, *Tales of Phantasia* (translated from Japanese to English by the group *DeJap*) being a reference for high standards in the romhacking community.

The outstanding quality of the translation for this game is to be found — apart from using good translators — in the fact that romhackers are very familiar with the games they translate, so they have context during the translation of the script. Unfortunately, this is not the situation for most professional translators, as they do not usually have access to the game itself (Bernal Merino, 2007, p. 2). Currently, only large companies, such as *Nintendo* or *Square Enix*, appear to hire translators to work in-house so they can play the game and have control over the translation from the beginning to the end of the localisation process (Mangiron, 2006, p. 310).

Another characteristic of romhacking projects is that they do not need to comply with censorship rules that forbid foul language, sexual content or copyright issues with brand names, as translations are unofficial and therefore do not undergo a legal check. For instance, the *DeJap* romhackers' translation of the *Super Nintendo* game *Tales of Phantasia* contains a scene in which the language is appropriate only for adults, while in the *Game Boy Advanced* version, which was localised by Namco in collaboration with *Nintendo*, the language is much more adequate for all audiences (Figure 6). Furthermore, some locations in the *Super Nintendo* version include swastikas, which do not appear in the *Game Boy Advanced* version (Figure 7).



Figure 6. Dialogue differences between the unofficial (left) and official (right) translations of *Tales of Phantasia*

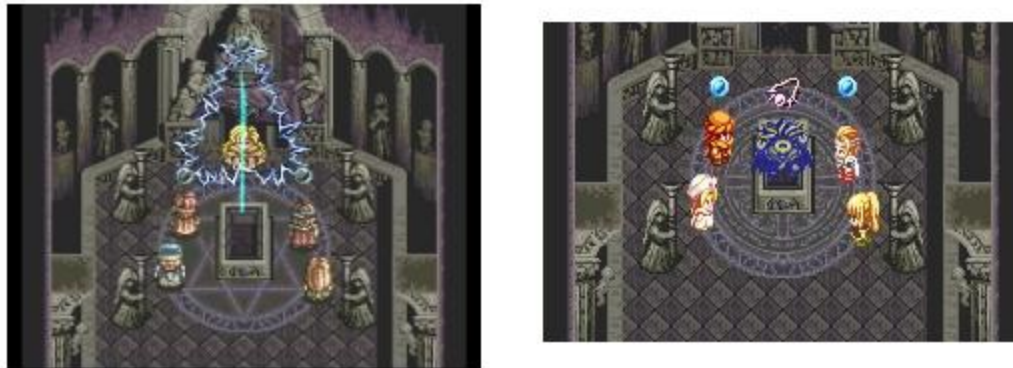


Figure 7. Graphic differences between the unofficial (left) and official (right) translations of *Tales of Phantasia*

As the game was translated unofficially by *DeJap* romhackers before the release of the Game Boy Advanced version, many fans criticised the official translation because of the differences, since the transcriptions of some character names were different from what they were used to seeing (as shown in Figure 7, “Klarth” appears as “Claus” in the official translation) and because the original Japanese voices were dubbed in English. This is the reason why some video games that were translated officially have been retranslated by romhackers in the interest of satisfying the fans. As Mangiron (2006) stated:

Fan translations have on occasion received higher praise than the official ones, such as the English translation of the Japanese RPG *Final Fantasy IV*, because fan translators, as gamers, know what the game should be like in order to appeal to other gamers like themselves. (p. 315)

Finally, some romhackers also rewrite the script of a video game to make it funny and give players the opportunity to play the video game from a different perspective. For instance, the Spanish romhacking group *Charnego Translations* made a parody of the famous video game

The Legend of Zelda: A Link's Awakening for the Game Boy Color. Apart from changing the graphics significantly from the original version (Figure 8), this group changed the entire story to make it hilarious, sexual and violent. This translation achieved so much popularity that the host of a Spanish well-known radio programme on video games called *Game Over* (2001) interviewed the authors themselves.



Figure 8. Comparison between the original (left) and parody (right) versions of *The Legend of Zelda: A Link's Awakening*

The Legality of Romhacking

As commercial ROMs contain copyrighted material, it is actually illegal to manipulate a ROM in order to translate it, according to the *Digital Millennium Copyright Act* (1998). However, game publishers do not usually take legal actions against romhackers because they normally translate old video games, which does not impact sales since these games become unavailable. Furthermore, this is a way to promote old video games within one of today's popular sagas, such as *Final Fantasy*.

Nevertheless, Pettus (1999, p. 88) described a case that dates back to 1999 in which the lawyers of a Japanese company called *ASCII* sent an e-mail to *KanjiHack*, a group of romhackers who were translating a computer game called *RPG Maker 95*, warning them to cancel the project

to avoid facing legal consequences. A more recent event took place in May of 2009; while it was not a translation project but a complete new game, the Japanese publisher *Square Enix* sent a cease-and-desist letter (Legal Department, 2009) to a group of fans who were creating a side story of the video game *Chrono Trigger* by using software used for romhacking projects. As they had not released their work to the public yet, they decided to cancel the project to comply with *Square Enix*'s demands.

Despite the fact that romhacking projects are carried out by and for fans with no profit motive in mind, they are considered illegal under current international laws, so romhackers must stop their work, if the copyright owners require this, to avoid facing legal action.

Conclusion

Video game localisation is different from any other type of translation due to the fact that end users are actually playing (that is, enjoying) when they read the translated texts. Therefore, as Mangiron and O'Hagan (2006) put it, "the main priority of game localisation is to preserve the 'look and feel' of the original" (p. 14). This is the reason why video game translation is often seen as "transcreation" (Mangiron & O'Hagan, p. 11), as translators opt for a domesticating approach in a Venutian sense (Venuti, 1995) so that players can amuse themselves without realising they are playing a translated game. Thus, translators must familiarise themselves with the video game subculture, and ideally, they should be gamers who know best the expectations of the target audience.

Thanks to romhacking, many gamers have had the opportunity to play video games which were never localised to their mother tongue. When this obscure process of video game localisation became easier with more user-friendly tools and new documentation, some gamers with good language skills even were able to translate by themselves the games they enjoyed in

their childhood. At no time did they think they were working for free so others could play in their mother language, as they were simply enjoying what they were doing.

Romhackers are rarely qualified translators, but they have the advantage of having a thorough knowledge of the game domain. Therefore, using the skills they develop, they find it easier to get involved with the video game industry. As Mangiron (2006) concluded, “Sometimes job specifications for localisation specialists do not require previous experience in this field, only some previous translation experience, experience as a gamer, and knowledge of the gaming world” (p. 315).

In this sense, participating in a romhacking project can be seen even as an educational aid for those seeking to work in the video game industry. In fact, such projects are a way for them to start familiarising themselves with common translation issues in video game localisation, such as the use of variables, text limitations within menus, dealing with special tags to insert pauses in the text, or other game specific issues. Further research is required to see if romhacking might help to fill the gap between academia and business, since to date, only a very limited number of universities offer video game localisation in their curricula.

References

- Bernal Merino, M. (2007). Challenges in the translation of video games. *Tradumàtica*, 5. Retrieved November 22, 2009, from <http://webs2002.uab.es/tradumatica/revista/num5/articles/02/02.pdf>
- Bernal Merino, M. (2006). On the translation of video games. *The Journal of Specialised Translation*, 6, 22-36. Retrieved June 7, 2009, from http://www.jostrans.org/issue06/art_bernal.pdf
- Burnham, V. (2001). *Supercade: A visual history of the videogame age*. Cambridge, MA: MIT Press.
- Chandler, H. (2006). An Interview with Heather Chandler, author of *The Game Localization Handbook*. Retrieved November 22, 2009, from http://www.bytelevel.com/global/game_globalization.html
- DeJap* (2003). Retrieved November 22, 2009, from <http://dejap.eludevisibility.org>
- Díaz Cintas, J. & Muñoz Sánchez, P. (2006). Fansubs: Audiovisual translation in an amateur environment. *The Journal of Specialised Translation*, 6, 37-52. Retrieved June 7, 2009, from http://www.jostrans.org/issue06/art_diaz_munoz.pdf
- Digital millennium copyright act, H.R. 2281, 105th Congress*. (1998). Retrieved June 20, 2009, from <http://thomas.loc.gov/cgi-bin/query/z?c105:H.R.2281>:
- Esposito, N. (2004). *Émulation et conservation du patrimoine culturel lié aux jeux video* [Emulation and curation of the video game cultural heritage]: *Proceedings of ICHIM 04 (Digital Culture and Heritage)*, Berlin. Retrieved June 8, 2009, from http://www.archimuse.com/publishing/ichim04/3391_Esposito.pdf
- Esselink, B. (2002). Localization engineering: The dream job? *Tradumàtica*, 1. Retrieved June 8, 2009, from <http://www.fti.uab.es/tradumatica/revista/articles/besselink/besselink.PDF>

Game over [Interview with two Spanish romhackers from the group *Charnego Translations*].

(2001). Retrieved June 17, 2009, from

<http://charnego.frikipedia.es/downloads/mp3/entrevista1.mp3>

Gieske, J. (2002). *Avoid missing ball for high score* (Master's thesis, The University of Salford,

2002). Retrieved June 9, 2009, from

<http://www.digitalgamearchive.org/data/news/gieske.pdf>

hack. (2009). In *Merriam-Webster Online Dictionary*. Retrieved June 28, 2009, from

<http://www.merriam-webster.com/dictionary/hack>

Legal Department. *Cease-and-desist letter to Chrono Compendium, Crimson Echoes from*

Square Enix (2009). Retrieved June 22, 2009, from

<http://www.chronocompendium.com/CEOrder.PDF>

Mangiron, C. (2006). Video games localisation: Posing new challenges to the translator.

Perspectives: Studies in Translatology, 14(4), 306-317.

Mangiron, C. & O'Hagan, M. (2006). Game localisation: Unleashing imagination with

“Restricted” Translation. *The Journal of Specialised Translation*, 6, 10-21. Retrieved June

24, 2009, from http://www.jostrans.org/issue06/art_ohagan.pdf

Muñoz Sánchez, P. (2007). *Manual de traducción de videojuegos: el fascinante mundo del ROM*

hacking [A handbook on video game localisation: the fascinating world of romhacking].

Retrieved June 10, 2009, from

http://www.lulu.com/items/volume_62/1098000/1098021/4/print/manualromhacking.pdf

Pettus, S. (1999). *Emulation: Right or wrong?* Retrieved June 21, 2009, from

<http://www.overclocked.org/emufaq/EmuFAQ.pdf>

Poole, S. (2000). *Trigger happy: Videogames and the entertainment revolution*. New York:

Arcade Publishing.

Pablo Muñoz Sánchez

Romhacking.net. (2009). Retrieved June 11, 2009, from <http://www.romhacking.net>

Sellers, J. (2001). *Arcade fever: The fan's guide to the golden age of video games*. Philadelphia: Running Press.

Square Enix. (2009, February 12). *Results briefing session. Nine-months period of the fiscal year ending March 31, 2009*. Retrieved June 8, 2009, from the Square Enix Web site: http://www.square-enix.com/eng/pdf/news/20090216_01en.pdf

Venuti, L. (1995). *The translator's invisibility*. New York and London: Routledge.