

KARINA GONÇALVES DE SOUZA DE OLIVEIRA

Loanword adaptation in Esperanto

Abstrakt (Adaptacja zapożyczeń w języku esperanto). Niniejsze badania zostały przeprowadzone celem zbadania kierunków fonologicznych, z których adaptowane są przez język esperanto nowe rdzenie wyrazowe. Poszczególne wyrazy wybrano spośród następujących czasopism: *Kontakto* – oficjalny magazyn Światowej Esperanckiej Organizacji Młodzieżowej (Tutmonda Esperantista Junulara Organizo – TEJO), który po raz pierwszy wydano w 1963 roku i który ma swoich abonentów w ponad 90 krajach świata, oraz *Esperanto* – oficjalny magazyn Światowego Związku Esperantystów (Universala Esperanta Asocio – UEA), który po raz pierwszy został wydany w 1905 roku i który ma swoich czytelników w 115 krajach na świecie, a także listy terminologii technicznej (Nevelsteen 2012) oraz wyrazów nie zarejestrowanych jeszcze w słownikach, lecz opublikowanych na liście na blogu <<http://vortaroblogo.blogspot.com.br/2009/09/nepivajvortoj-i.html>>. Wyrazy pochodzą z 13 języków, tj. arabskiego, chińskiego, francuskiego, angielskiego, japońskiego, komi, koreańskiego, portugalskiego, rosyjskiego, hiszpańskiego, tureckiego, sanskrytu oraz suahili. Podstawą teoretyczną niniejszej analizy jest Fonologia Zapożyczeń (Loanword Phonology); głównie Calabrese & Wetzels (2009), Vendelin & Peperkamp (2006), Paradis (1988), Kang (2011), Friesner (2009), Menezes (2013), Chang (2008), Kenstowicz & Suchato (2006) oraz Roth (1980). Analiza korpusu wykazała, że wyrazy mogą być dostosowywane do języka esperanto zarówno poprzez ich formy fonetyczne, jak i poprzez oryginalną ortograficzną formę zapisu bezpośrednio z danego języka. Co więcej, zaobserwowano, że samogłoski długie zostały w przeważającej mierze zaadaptowane jako samogłoski proste, zaś niektóre wyrazy mogą występować w dwóch wersjach synchronicznych.

Abstract. This research investigated the phonological directions by which new roots are incorporated into Esperanto. Words were selected from the following magazines: *Kontakto*, the official magazine of the *Tutmonda Esperantista Junulara Organizo* (TEJO – World Esperanto Youth Organization), which was first published in 1963 and has subscribers in over 90 countries, and *Esperanto*, the official magazine of the *Universala Esperanto-Asocio* (UEA – Esperanto Universal Association), which was first released in 1905 and has readers in 115 countries, in addition to a technological terminology list (Nevelsteen, 2012) and to words not quoted in dictionaries but published in a list on the blog <<http://vortaroblogo.blogspot.com.br/2009/09/nepivajvortoj-i.html>>. Words were collected from 13 different languages: Arabic, Chinese, French, English, Japanese, Komi, Korean, Portuguese, Russian, Spanish, Turkish, Sanskrit and Swahili.

The theoretical basis that guided this analysis was *Loanword Phonology*, mainly the works of Calabrese & Wetzels (2009), Vendelin & Peperkamp (2006), Paradis (1988), Kang (2011), Friesner (2009), Menezes (2013), Chang (2008), Kenstowicz & Suchato (2006) and Roth (1980). An analysis of the corpus showed that words can be adapted by their phonetic form as well as by their root's orthographic form from the original language. Furthermore, we observed that long vowels were, for the most part, adapted as simple vowels; and some words are present in two synchronic variations.

1. Introduction

Esperanto is a planned language that follows the same patterns of variation and changes that non-planned languages experience, because it has evolved to become a natural language. This paper shows the results of research that explored how new words (roots) are incorporated into Esperanto in order to try to understand better the phonology of the language. Since it is a planned language, Esperanto is a little different from other languages.

Zamenhof's words show the way in which his work differs from that of a linguist trying to codify a language. For a linguist struggling with an outbreak of dialects, the task would be to produce a standardized written form for every word. In the opposite direction, for the conlanger with a view to producing an international auxiliary language in the age of writing, the process would begin with a written word and the effort applied to the need for restricting oral variation. Natural languages are first spoken, and only a small number of them have achieved a written form. Artificial languages grow the other way round (Dols Salas 2012: 38).

Esperanto, then, had its orthography planned before it could manifest the phonetic reality of the variants. Today, it is a language spoken around the world, with different accents, as any other language. The phonology of the language, however, is coherent and has its own rules. The study of new words in the language was a way to try to observe better how these phonological rules work.

2. Methodology for the selection of words

To select loanwords in the language, we decided to collect roots from two important magazines in the Esperanto community (*Kontakto*, published since 1963 and *Esperanto*, published since 1905), a book (Nevelsteen, 2012), which is a list of terminology in the semantic field of technology and words listed as “not yet in the main dictionary” of the language from John C. Wells's blog (the Uni-

versity College London). The words were included in the corpus if they were not in the official dictionary of Esperanto, named *PIV* (*Plena Ilustrita Vortaro*) and if examples of usage could be found on the internet (to confirm that the community really uses the words).

The pdf files of the magazines (editions between 2004 and 2014) were collected from the official website of UEA (<http://www.uea.org/>), transformed in .txt files and entered into a computational program that used the language *Python* to run a list of all the words in the magazines. By doing this, we obtained a list of the words from 183 issues of these two magazines. After that, the list was verified to check all the new possible words (the entire list of new words is available in Oliveira 2016).

After selecting the loanwords from that list, we organized them based on the different languages they came from (13, in total, despite the fact that most of them came from English): Arabic (04), Chinese (02), French (01), English (37), Japanese (10), Komi (01), Korean (01), Portuguese (03), Russian (01), Spanish (02), Turkish (02), Sanskrit (01) and Swahili (01).

3. Data and Analysis

In order to analyze the corpus, we have utilized the field referred to as *Loanword phonology*. When different phonological systems have contact with each other, every language has its own way to adapt new words. According to Calabrese & Wetzels (2009: 01):

Speakers borrow words from other languages to fill gaps in their own lexical inventory. The reasons for such lexical gaps vary greatly: cultural innovation may introduce objects or actions that do not have a name in the native language; native words may be perceived as non-prestigious; names of foreign cities, institutions, and political figures which were once unknown may have entered the public eye; new words may be introduced for play, etc.

So, when a new word appears in a language, if it has some sound that is not part of the phonological system of the language that it is entering, adjustments have to be made, and those adjustments are made in different ways. For example, the phoneme [y], from French, is adapted as [u], [i] or even [ju] by different languages (Kang 2011: 08). Research points to the fact that both phonological and phonetic factors (as well as morphological, semantic and orthographic ones) are important for the results obtained after various adaptations are applied (Kang 2011: 12).

We have observed that consonants that are part of the phonological system of Esperanto were adapted with no changes, such as **blogo** ['bloɡo] < *blog* ['blɒɡ], from English, for example, where there is barely any modification. The phonemes

[b], [l] and [g], present in both languages, were kept as the same in the adaptation. The same occurs with the word **Salso** ['salso], from the Spanish word *Salsa* ['salsa] (a kind of dance) and **tabasami** [taba'sami], from Swahili, *tabasamu* [taba'samu] (“to smile”).

However, when a consonant is not part of the phonology of Esperanto, the adaptation varies. The phonemes [tɛ] and [tɛ^h] are adapted as [ʃ] in the words **vejĉio** [vej'ʃio] < *wéiqí* [wejtɛ^hi] (from Chinese “goo game”), **peĉakuĉo** [peʃa'kuʃo] < *pechakucha* [petɛakutea] (from Japanese, a specific kind of public presentation) or **ĉumo** > [ʃumo] < чум [tɛum] (from Komi – a language spoken in Russia –, a kind of house used by nomadic populations in Siberia). In these cases, the orthography of the transliterations may have influenced the adaptations.

One of the patterns found in the data was the adaptation of the word by *adaptation as perception*, as explained by Kang (2011: 05).

The following are those words that follow these patterns: **ĉaĉao** [ʃa'ʃao] < *cha-cha-cha* [ʃaʃa'ʃa] (from Spanish, a kind of dance); **androjdo** [an'drojdo] < *android* ['ændrɔjd], **bajto** ['bajto] < *byte* ['bajt], **dizajno** [di'zajno] < *design* [di'zajn], **fejsbuko** [fejs'buko] < *facebook* ['fejsbʊk], **gigabajto** [giga'bajto] < *gigabyte* ['gigəbajt], **giko** ['giko] < *geek* [gi:k], **ĝavo** ['dʒavo] < *java* ['dʒavə], **megabajto** [mega'bajto] < *megabyte* [megəbajt], **Skajpo** ['skajpo] < *Skype* ['skajp], **splajno** ['splajno] < *spline* [splajn], **terabajto** [tera'bajto] < *terabyte* [terəbajt] (from English); **kaŝaso** [ka'ʃaso] < *cachaça* [ka'ʃasɐ] (an alcoholic beverage), **pandero** [pan'dero] < *pandeiro* [pɛn'derʊ] (a musical instrument, from Portuguese).

This approach breaks away from the assumption that the input to the production grammar in loanword adaptation faithfully retains the phonetic and/or phonological structure of the source language input (cf. Jacobs & Gussenhoven 2000, La Charité & Paradis 2005). This view provides a solution to many puzzling adaptations, such as *unnecessary repair* or *divergent repair*, where the adaptation pattern seems to contradict the production grammar of the borrowing language (Kang 2011: 05).

According to Kang and other researchers in the field, this is one of the many possible phases of the adaptations, but others say that the adaptations occur, primarily, during the perception of the way the words are spoken (Kang, 2011 : 06).

In the words of our corpus, we can observe that there are elements that indicate the adaptation closer to the phonetic variance of the word, despite its written form. In the words **bajto** ['bajto] < *byte* ['bajt], **dizajno** [di'zajno] < *design* [di'zajn], **fejsbuko** [fejs'buko] < *facebook* ['fejsbʊk], **gigabajto** [giga'bajto] < *gigabyte* ['gigəbajt], **megabajto** [mega'bajto] < *megabyte* [megəbajt], **Skajpo** ['skajpo] < *Skype* ['skajp], **splajno** ['splajno] < *spline* [splajn] and **terabajto** [tera'bajto] < *terabyte* [terəbajt] (from English) there is, in the Esperanto words, a diphthong that probably comes from the phonetic form of the English words, because in the written form it is represented only by one letter.

The word *android* [ˈændrɔjd] has a sequence of letters (<oi>) that creates two variants in Esperanto: **androjdo** [anˈdrojdo] and **androido** [androˈido]. The first looks like the phonetic form /ɔj/ was considered to create the form /oj/ in Esperanto, but in the second case a hiatus arises, probably because of the orthographic form. The same phenomenon occurs with the word *java* [ˈdʒavə], that has two forms in Esperanto: **ĝavo** [ˈdʒavo] (phonetic similarity inter the first sound) and **javo** [ˈjavo] (orthographic similarity with the first letter).

The Portuguese word *pandeiro* [pənˈdeɾo] was adapted as **pandero** [panˈdeɾo], which shows that the phonetic form was taken, because of the absence of the letter <i> after <e> (which appears in the orthographic form but not in the phonetic one).

As for the word *geek* [gi:k], adapted as **giko** [ˈgiko], we should consider, beyond the phonetic influence, the fact that the word *geko* already exists in Esperanto (meaning *lizard/gecko*) so, the phonetic form was chosen to avoid the creation of a homonym.

On the other hand, the influence of the orthography can be seen in some words too. Vendelin and Peperkamp (2006) say that the speakers can adapt the words as they were reading it in the language they adapt them from. That way the orthography tends to be maintained, despite the phonetic form. This is an important issue for loanwords, as Wüster (1979) observed in the field of the terminology as well. This author says that the written form is a better option for the internationalization of terms. An example of this is the word “psychology” (in English) and the word *Psychologie* (in German), which has more in common in the orthography than in the phonetic pronunciation.

Consider the following words: **androido** [anˈdroˈido] < *android* [ˈændrɔjd], **animacio** [animaˈtsio] < *animation* [æniˈmeɪʃən], **ciberspaco** [tsibersˈpatso] and **kiberspaco** [kibersˈpatso] < *cyberspace* [ˈsajbəspejs], **Emakso** [eˈmakso] < *Emacs* [iˈmæk], **javo** [ˈjavo] < *java* [ˈdʒavə], **haloveno** [haloˈveno] < *Halloween* [hæləwˈi:n], **iPodo** [iˈpodo] < *iPod* [ˈajpɒd], **kibernetiko** [kiberneˈtiko] < *cybernetics* [sajbəˈnetiks], **Makintoŝo** [makinˈtoʃo] < *Macintosh* [ˈmækɪntɒʃ], **skiflo** [ˈskiflo] < *skiffle* [ˈskɪfl̩], **Unikso** [uˈnikso] < *Unix* [juːnɪks] and **vifio** [viˈfio] < *wi-fi* [ˈwajfaj].

The absence of a diphthong in the words **android** [anˈdroˈido] < *android* [ˈændrɔjd] ([o.i] < [ɔj]), **animacio** [animaˈtsio] < *animation* [æniˈmeɪʃən] ([a] < [ej]), **haloveno** [haloˈveno] < *Halloween* [hæləwˈi:n] ([o] < [əw]), **iPodo** [iˈpodo] < *iPod* [ˈajpɒd] ([i] < [aj]), **Unikso** [uˈnikso] < *Unix* [juːnɪks] ([u] < [juː]) and **vifio** [viˈfio] < *wi-fi* [ˈwajfaj] ([i] < [aj]) shows the influence of the orthography, because the diphthongs, present in the phonetic forms of the words, have disappeared.

There is yet another word that can be analyzed according to that pattern: **ŭonbulismo** [wonbuˈlismo] < *Won Buddhism* [wʌnbul(gjo)] (from Korean). The sequence <ŭo> does not appear in other words in Esperanto, but it was allowed

in this case, probably because of the orthography. The pattern called *differential importation* can explain that.

Importation refers to a situation where a structure not attested in native phonology is exceptionally allowed in loanwords. While such importation in and of itself is not a problem from a learnability perspective, the fact that only certain structures, but not others, are imported requires an explanation (Holden 1976, Itô & Mester 1995, 1999, 2001, Davidson & Noyer 1997, Broselow 2009). Given foreign input with two types of novel structures which are both equally unattested in the native data, why is one structure readily allowed into the language but not the other? For example, in Hawaiian, the fully nativized form of the English word *truck* is [kə'lakə]. Also possible is a “less Hawaiian” variant [tə'lakə], where English /t/ remains unadapted. But the variant *[ˈkrakə], where the complex onset is retained, but /t/ is adapted as /k/, is judged to be impossible (Adler 2006). In other words, the restriction against /t/ is more easily relaxed than the restriction against an onset cluster. In Russian, the requirement that only palatalized consonants occur before /e/ is often violated in adaptation, but the process of reducing unstressed /o/ and /e/ is more likely to be upheld (Holden 1976 in Kang 2011: 03).

To be able to verify which structures can be allowed in Esperanto and which ones cannot, the analysis of more data is needed.

Regarding vowels, there were many cases of long vowels that were adapted in Esperanto as short vowels, because there are no long vowels in the Esperanto system: **ĝihado** [dʒi'hado] < *jihād* [dʒi'ha:d] (from Arabic); **giko** ['giko] < *geek* [gi:k], **haloveno** [halo'veno] < *Halloween* [hæləw'i:n], **jutubo** [ju'tubo] < *YouTube* ['ju:tu:b] and **Vikipedio** [vikipe'dio] < *Wikipedia* [wɪki'pi:diə] (from English); **sudoko** [su'doko] < *suudoku* [su:doku], **sumoo** [su'moo] < *sumoo* [sumo:], **ŝogio** [ŝo'gio] < *shoogi* [eɔ:gi] and **tofuo** [to'fuo] < *tōfu* [to:ɸu] (from Japanese). In most of these cases, the long vowel was adapted as a short vowel of the same quality, when the vowel was present in the system. However, the Japanese phoneme /u/ was adapted as /u/, that is, a non-rounded vowel was adapted as rounded, because it was the most similar in the system that adapted it.

As regards the word **tofuo** [to'fuo] < *tōfu* [to:ɸu] (from Japanese), it's important to consider that there are, also, other possible forms: *toŭfuo* [tow'fuo] kaj *tohuo* [to'huo] (Wennergren 2005). This author affirms that *tofuo* or *tohuo* are the optimal variants, but *toŭfuo* is present in PIV (the official dictionary of the language). The diphthong <ou> could be a way of preserve the long vowel of the source word in Japanese, by maintaining a nucleus of the syllable with two positions filled. This can be explained by Kang (2011) who summarizes one of the patterns of the Loanword Theory as “*The too-many-solutions problem*”: there is always more than one way to repair some structure that is not allowed in the language, but speakers usually use the same one.

For example, Hawaiian does not have a voiced stop /b/, and thus, as it is unattested in native phonology by definition, there is no direct evidence from the native phonology

as to how such an illicit segment should be repaired. Yet English /b/ is systematically adapted as /p/ (*boulder* → [polu'ka:]) and not /m/, /w/ or any other segment of the Hawaiian inventory (Adler 2006 in Kang 2011: 02).

Another case that allows several possibilities is the syllabic pattern.

For example, when an onset cluster (C_1C_2V) is borrowed into a language which bans complex onsets, the structural requirements of the native language can be satisfied by the deletion of a consonant ($> C_2V$ or C_1V), an option found in French loanwords in Vietnamese (*crème* → [kem]), or by the epenthesis of a vowel in front of the cluster ($> vC_1C_2V$) or inside the cluster ($> C_1vC_2V$). The epenthesis repair is found in Japanese (*Christmas* → [kurisumasu]) and Hawaiian (*Christmas* → [kalikimaki]), among other languages, and these examples also illustrate some of the different possibilities in the quality of the epenthetic vowel (Broselow 2006 in Kang 2011: 02).

So, it seems that even if the possibility exists of adapting a long vowel as either a short one or as a diphthong, speakers prefer the first option.

Another important pattern of the adaptations is the effect of factors that are non-grammatical.

[...] it has been proposed that sociolinguistic or grammar-external factors affect the pattern of (non-)adaptation, especially where aspects of loanword phonology are underdetermined by grammatical factors. First of all, the rate of importation has been shown to positively correlate with the level of bilingualism in the community (Haugen 1950, Paradis & La Charité 1997, 2008, 2009, Heffernan 2007, Friesner 2009a in Kang 2011: 7).

Because of the fact that Esperanto is a planned language, the community regularly discusses the possibility of new words (through social media or during Esperanto meetings). What remains after the discussions (or despite them) is the real use of the words, of course. For example, there are two adaptations of the English word *facebook*: *vizaĝlibro* and *fejsbuko*. The first word is a semantic translation of the English words *face* (*vizaĝo*) and *book* (*libro*). Many people defend the idea that it is not necessary to introduce a new root in the language to the word *facebook*, because the language already has elements that can be combined to form a good word to that. However, it seems that the variant *fejsbuko* is, in fact, more commonly used.

Finally, there are some words that have two variants in Esperanto, that are the results of different strategies of phonological adaptation:

1. **tajĝiĉjŭano** [tajdʒi'ʃjwano] or **tajĝio** ([tajdʒi'o]), from the Chinese word *tàijíquán* [tʰajtɕiteʰwan],
2. **androjdo** [an'drojdo] or **androido** [an'dro'ido], from the English word *android* ['ændrɔjd],
3. **ciberspaco** [tsiberspato] or **kiberspaco** [kiberspato], from the English word *cyberspace* ['sajbəspejs],

4. **ĝavo** ['dʒavo] or **javo** ['javo], from the English word *java* ['dʒavə],
5. **karaoko** [kara'oko] or **karaokeo** [karao'keo], from the Japanese word *karaoke* [karaoke].

Probably one of these two variants will prevail over the other in the future, but for now they are both used by the community.

Because of the fact that Esperanto is a planned language, some words are suggested, intentionally, by the users. One such word is *tabasami*, which means “to smile”. The word “rideti” already exists in Esperanto, but one speaker thought it was not good enough to express the meaning, so, he suggested the Swahili word. It’s adoption, or course, depends on whether other speakers use it or not.

4. Conclusions

The data show that the scope of the adaptation of loanwords in Esperanto is large and loanwords occur often, as in any other language. We have shown some adaptations of consonants and vowels, as well as other kind of adaptations. Further research is needed to better understand the patterns of the adaptations.

Bibliography

- Calabrese, A. & Wetzels, W. L. (eds) 2009: *Loan phonology*, vol. 307. Amsterdam/Philadelphia: John Benjamins Publishing.
- Chang, C. B. 2008: Phonetics vs. phonology in loanword adaptation: Revisiting the role of the bilingual. *Annual Meeting of the Berkeley Linguistics Society*, vol. 34(1): 43-54. http://linguistics.berkeley.edu/phonlab/annual_report/documents/2008/chang_bls34.pdf
- Dols Salas, N. 2012: Phonology and morphology and the limits of freedom in an artificial language. *Język. Komunikacja. Informacja*. tom 7. <http://hdl.handle.net/10593/9809>; <http://jki.amu.edu.pl>.
- Friesner, M. L. 2009: The adaptation of romanian loanwords from Turkish and French. In: *Loan phonology*, vol. 307: 115-129.
- Kang, Y. 2011: Loanword phonology. *The Blackwell companion to phonology*, vol. 4: 1003-1026.
- Kenstowicz, M. & Suchato, A. 2006: Issues in loanword adaptation: A case study from Thai. *Lingua*, vol. 116(7): 921-949. http://web.mit.edu/~linguistics/people/faculty/kenstowicz/loanword_adaptation.pdf
- Menezes, M. B. 2013: *A adaptação de empréstimos recentes no papiamentu moderno*. Master Dissertation. University of São Paulo, São Paulo.
- Nevelsteen, Y. 2012: *Komputeko: prikomputila terminoteko*. Partizánske: Espero.
- Oliveira, K. G. S. 2016: *Adaptação de Empréstimos em Esperanto*. Master Dissertation. University of São Paulo, São Paulo.
- Paradis, C. 1988: Towards a theory of constraint violations. *McGill working papers in linguistics*, vol. 5(1): 1-43.

- Roth, W. 1980: O empréstimo como problema da lingüística comparada. *ALFA: Revista de Linguística*, volume 24: 157-177. <http://seer.fclar.unesp.br/alfa/article/view/3620/3389>.
- Vendelin, I. & Peperkamp, S. 2006: The influence of orthography on loanword adaptations. *Lingua*, vol. 116(7): 996-1007. [http://www.lscp.net/persons/peperkamp/Vendelin_Peperkamp_\(2006\)_The_influence_of_orthography_on_loanword_adaptations.pdf](http://www.lscp.net/persons/peperkamp/Vendelin_Peperkamp_(2006)_The_influence_of_orthography_on_loanword_adaptations.pdf)
- Wennergren, B. 2005: *Plena Manlibro de Esperanta Gramatiko*. Sacramento, California: Esperanto-Ligo por Norda Ameriko. <http://bertilow.com/pmeg/>.
- Wüster, E. 1979: *Introduction to the General Theory of Terminology and Terminological Lexicography*. Viena: Springer.