# A CLASSIFICATION OF BORROWINGS: OBSERVATIONS FROM ROMANIAN/ENGLISH CONTACT

Arina GREAVU
"Lucian Blaga" University, Sibiu
arinagreavu@yahoo.com

**Abstract:**
The classification of language contact phenomena has always been an important concern among researchers in the field. In particular, the term borrowing has received different definitions from different writers, covering a wide spectrum of words, from pure loanwords to hybrid loans and semantic extensions. This paper presents one of the most influential taxonomies of borrowings in the literature, and analyzes the way in which the various categories proposed in this taxonomy apply to the Romanian/English contact situation. English borrowings selected from a corpus of journalistic prose and from the specialized literature are used to illustrate the theoretical discussion.

**Key words:**
Borrowing, loanword, loanblend, loanshift.

### 1. Introduction

The classification of words borrowed from one language into another was one of the first aspects of their study to engage the attention of researchers in the field. This special interest in the categorization of borrowings was motivated by the fact that, in the early stages of language contact study, the emphasis was mainly on the products of borrowing rather than on the process in itself (J. Treffers-Daller, 2000, p. 2). At the same time, it was a consequence of the large array of possible combinational patterns between native and foreign material, which in turn reflected the highly complex character of the borrowing process. These realities formed the backdrop against which various proposals for classification emerged, some of which are still in use today. Thus, Winford (2003, p. 42) reports on attempts to classify borrowings as early as the end of the 19th and the beginning of the 20th century, for example by Paul (1886), Seiler (1907-1913), Eugene Kaufman (1939), and Betz (1949). However, one of the most comprehensive taxonomies of borrowings was developed by Haugen (1950, 1956) based on the analysis of the speech of Norwegian immigrants into the United States.

The main tenet of Haugen's theoretical framework is that borrowing results from the joint action of two mechanisms, importation and substitution. Importation occurs when a foreign word is reproduced in a language so that it can be unmistakably tracked back to the model. Substitution, on the other hand, involves the replacement of some morphemes in the source language word by recipient

language ones, in an attempt to integrate it into the structures of this language. Haugen is of the opinion that "*speakers are in a rough way carrying on an operation of linguistic comparison between the two languages*"[1], with the following result:

"*Any likeness between them is importation, while any difference between them is regarded as substitution of native material. Substitution means that the imitation of the foreign model is less than perfect, but it also means that it has become more familiar to those who speak the native language.*"[2]

Depending on the ways in which importation and substitution combine in the borrowing process, the outcomes of this process can range on a form-meaning continuum from foreign forms being borrowed together with their meanings, to meanings entering a language on their own. In broad lines, Haugen's taxonomy of borrowing includes two main categories – loanwords and loanshifts – each of these containing other subcategories. Thus, loanwords are divided into pure loanwords (unassimilated, partly assimilated and wholly assimilated) and loanblends (derivative and compound). Loanshifts, in their turn, comprise extensions or semantic loans, and creations.

In the following sections of this paper we are going to illustrate these separate classes of borrowings with examples from a corpus of Romanian journalistic prose (i.e. one year of the economic publication *Capital*), and from the Romanian literature on the topic.

### 2. Corpus[3] and methodology of research

The source of the corpus was the economic magazine *Capital* on CD-ROM, consisting of PDF files[4]. Following a process of English words identification and filtering so as to eliminate Romanian homographs and proper nouns[5], we arrived at the final amount of data: 1,442 borrowed types occurring in a total of 20,534 tokens. All these words were tagged according to the formation process from which they resulted, thus allowing for conclusions regarding the numerical

---

[1] Einar HAUGEN, 1956, *Bilingualism in Americas: A Bibliography and Research Guide*, University of Alabama Press, p. 388.

[2] .*Ibidem*.

[3] The annotated corpus was used as part of a PhD project on recent English borrowings in Romanian, conducted at the "Babeş-Bolyai" University of Cluj Napoca.

[4] This raw data underwent a series of processing procedures, i.e. Optical Character Recognition, sentence splitting, tokenization and part-of-speech tagging and lemmatization. All these processing tasks were performed by Eckhard Bick (researcher) and Tino Didriksen (student assistant), from the Institute for Language and Communication (ISK) at the University of Southern Denmark. The tagging was done using the MSD tagger developed by the Research Institute for Artificial Intelligence of The Romanian Academy, under Professor Dan Tufiş' supervision. The pos-tagged corpus is available at http://corp.hum.sdu.dk/cqp.ro.html.

[5] Our method for the identification and filtering of English borrowings partly follows the method used by Onysko (2007) in his corpus study of Anglicisms in German.

representation of pure loanwords and of loanblends in the total of borrowings, as well as the relations these two classes hold with each other. The main focus of analysis in the present paper is constituted by the first category of borrowings in Haugen's taxonomy (i.e. loanwords), which is discussed both qualitatively and quantitatively. Some brief considerations regarding loanshifts will also be present, although the examples used to illustrate this class of borrowings (mainly from the Romanian literature on the topic) are not analysed quantitatively, the discussion being conducted mainly in descriptive, qualitative terms.

### 3. Presentation of results and discussion
### 3.1. Loanwords

The first important category in Haugen's classification of borrowings is that of loanwords. Loanwords "*show morphemic importation without substitution*"[6], but with some degree of phonological substitution. In other words, a language borrows both the form and the meaning of a foreign word, which may undergo a process of phonetic integration into the structures of the recipient language. According to their level of adaptation, loanwords can be described as unassimilated, partly assimilated and wholly assimilated.

Some of the assimilated loanwords that have been identified in the studied corpus, or have been reported by other researchers (Ciobanu 1996, Manolescu 1999) in their studies on Anglicisms in Romanian are: *administraţie, bancnotă, box, boiler, brec, budincă, canoe, cargo, casetă, cec, cent, chec, chicinetă, cocher, colocvial, cocteil, corner, crichet, cros, derbi, dischetă, doc, docher, duplex, electron, fan, fault, finiş, folclor, fotbal, golf, handicap, henţ, hipi, hol, iard, interviu, jeanşi, laburist, lider, lift, pasa, picnic, picup, pocher, pop, punci, recesiune, reporter, rugby, sandviş, scheci, scor, seif, slip, smoching, sport, star, start, stoc, stop, stres, şampon, şerif, şiling, şort, şut, tenis, test, tichet, tramvai, trenci, troleibuz, trust, video, volei, trening, laburist, televiziune, tehnologie.*

However, depending on a number of several factors (e.g. speakers' proficiency in the source language and attitude towards borrowing in general, the frequency with which the loanword is used and its age of existence in the borrowing language), phonetic integration may be slowed down or not take place at all. In this case, we are dealing with pure loanwords, or foreign words in which importation takes place in the absence of any substitution. The corpus of *Capital* 2005 contains approximately 850 English words (lemmas) which can be described as relatively unassimilated, the main marks for their inclusion in this category being their formal identity with the model they copy. These recent borrowings are used in a number of 1,339 types and 19,395 tokens, a situation which indicates a repetition rate of about 14 for each borrowed type. Examples of unassimilated loanwords in the studied corpus include: *advertising, advocacy, airbag, brand,*

---

[6] Einar HAUGEN, 1950, "The Analysis of Linguistic Borrowing", in: *Language*, 26 (2), p. 214.

97

*business, cash, card, dealer, email, leasing, software, rating, broker, job, notebook, futures, laptop, lobby, chart, charter, boom, wireless, leadership, outsourcing, player, hardware, showroom, weekend, coach, knowhow, flash, hobby, talk-show, roaming, target, all-inclusive, feedback, teambuilding, mouse, show, open-source, derby, download, outdoor, browser, shopping, offshore, outplacement, board, copywriter, desktop, research, banner, bearish, bullish, freelancer, brief, smartphone, bestseller, trendy, voucher, zoom, lowcost, dressing, e-tax, blockbuster, shipping, etc.*

The borderline between assimilated and unassimilated loanwords cannot be easily drawn, the idiosyncratic and variable character of the integration process making it possible for the same word to exist in a language at different stages of phonetic and morphosyntactic adaptation (C. Myers-Scotton, 1993). In this context, a special situation is constituted by those unadapted borrowings that double already integrated forms: *bungalow vs. bungalov, cocktail vs. cocteil, cricket cs. crichet, ski vs. schi, yacht vs. iaht, leader vs. lider, derby vs. derbi, inchi vs. inci*. The introduction of a foreign word into a language at different times and more than once, in spite of its existence in an already assimilated form, was described by Haugen (1956) as loanword "*re-borrowing*", being seen as the result of different, co-existing stages of bilingualism within a speech community whose members are becoming increasingly exposed to a foreign language. We believe that the present-day Romanian society constitutes a fertile ground for loanword re-borrowing, as its members, being more and more exposed to English as the international lingua franca of the contemporary world, are adopting words that were borrowed in the past and exist as established borrowings.

Similar studies regarding the impact of English on other European languages have shown that when such doublets are present, the more recent and modern word tends to replace the older, assimilated form. For example, Onysko[7] presents evidence which suggests that the integrated forms *klub* and *handikap* are being displaced in German by the more recent borrowings *club* and *handicap*, presumably as a result of some special psycho-social factors surrounding German/English contact.

Although the relationship between the words in such pairs will not be studied any further in the present paper, we tend to believe that these recent lexical entries will continue to strengthen their position in front of their older counterparts, the prestige English enjoys as the language of modernity, and the increasing level of English proficiency among Romanian speakers being factors which can considerably slow down the adaptation of words borrowed from this language. The tendency towards the etymological writing of Anglicisms on the English model was noticed as early as 1997 by Mioara Avram, who predicted that this

---

[7] Alexander ONYSKO, 2007, Anglicisms in German: Borrowing, Lexical Productivity, and Written Codeswitching, Berlin, New York: Walter de Gruyter, pp. 64-67.

phenomenon would continue[8]. Thus, as noticed by Haugen in his study of American/Norwegian speech, the sociolinguistic circumstances surrounding the contact situation seem to override time-depth when it comes to the integration of borrowings, a situation which makes the relationship between assimilated and unassimilated loans rather problematic.

### 3.2. Loanblends
The second important category of borrowings in Haugen's taxonomy consists of loanblends or hybrid loanwords. A loanblend results from both morphemic substitution and importation. In other words, some part of the form of a foreign word is borrowed, while some part is replaced with native material. According to the type of native morphemes used to replace the foreign ones (whole words or affixes), loanblends are divided into blended derivatives and blended compounds.

'Blended derivatives' occur when native derivational suffixes are substituted for the foreign. For example, in PaG (Pennsylvania German) *–ig* is often substituted for the English *–y*, as in *bassig* 'bossy', *fonnig* 'funny', *tricksig* 'tricky'[9], while in Romanian *-are* often replaces the English *–ing*, as in *targetare* 'targeting', *brandare* 'branding', *clonare* 'cloning'. Haugen believes that blended derivatives occur especially when there is some formal resemblance between the two affixes, as in his examples above. Other writers prefer to use the term 'adaptation' as a general label for the mechanism that generates blended derivatives. For example, Avram[10] believes that in all Anglicisms derived with a Romanian affix, this has resulted from the adaptation of an English affix, e.g. *-aj < –age*, *-are < –ing, etc*. Finally, some writers are of the opinion that blended derivatives result from language internal derivational processes, not from the replacement of one affix by another. For example, Winford claims that most loanblends appear when *"native (recipient language-RL) derivational processes are applied to previously imported words"*[11], being the result of the general process of loanword morphological adaptation to the recipient language.

The *Capital* 2005 corpus contains more than 60 blended derivative types, which occur in a total of approximately 600 tokens. This statistic includes only the English roots described as unassimilated in the previous paragraphs. However, it should be said that according to some constraints proposed in the language contact literature, a foreign word which combines with a native affix must be phonologically integrated. Most notably, Poplack's Free Morpheme Constraint

---

[8] Mioara AVRAM, 1997, *Anglicismele în limba română actuala*, Conferință prezentată la Academia Română. București: Editura Academiei Române, p. 14.
[9] E. Haugen, 1956, p. 399.
[10] Mioara AVRAM, 1997, p. 23.
[11] Donald WINFORD, 2003, *An Introduction to Contact Linguistics,* Blackwell Publishing, p. 44.

99

(1980) clearly states the impossibility of mixing between stems and bound morphemes, unless they carry the same language index. From this perspective, it can be argued that, for example, the word *brand* has a different phonological status when used independently and when used with the native suffix *–are*. While this may be true at a very fine level of analysis, in this paper we preferred to regard a word as belonging to the same class in Haugen's taxonomy, as long as it preserved the same spelling.

The most common Romanian suffix combining with English roots is *–are*. This is substituted for three English suffixes: *ing* (e.g. *printare<printing, brandare<branding, scanare<scanning, targetare<targeting, updatare<updating, upgradare<upgrading, marketare<marketing)*, *-ship* (e.g. *sponsorizare<sponsorship*), and *–ation* (e.g. *containerizare<containerization*). Other native suffixes used in blended derivatives are *–aj,* which replaces the English *age* (e.g. *brokeraj<brokerage*), *-ist*, which replaces the English *–er* (e.g. *retailist<retailer, lobbyist<lobbyer* but also *lobbyst*), and the verb forming *–iza* which is substituted for the English *–ize* (e.g. *computeriza<computerize, containeriza<containerize*).

The most productive Romanian prefix in terms of its ability to combine with English roots is *re-* (e.g. *rebrandare, resetare, rebrandata, rebrenduita*), followed by *hiper-* (e.g. *hiper-retail, hipermarket*) and *co-* (e.g. *cobrandata*), while the most frequently used one is *hiper-* (in *hipermarket),* with 131 tokens of occurrence. Although both *re-* and *co-* exist in English as well, we have chosen to regard them as native in the examples above due to the fact that the root they attach to is suffixed in Romanian, and therefore phonetically integrated according to Poplack's formal constraint. However, other possible cases of derivative loanblends were left out of the analysis as ambiguous (e.g. *sub-trend, teleshopping, multimedia, minicard, agribusiness*), in the absence of pronunciation cues being impossible to label the prefixes involved as either native or foreign. Without categorically excluding a Romanian extraction for these prefixes, we contend with Haugen that, when there is strong formal similarity between words *"it may be impossible to determine whether any transfer has taken place."*[12].

The same English root can be used with a Romanian affix in some cases, and with an English one in others. We believe that such a situation indicates an ongoing process of adaptation of the derived borrowings in question, a scenario supported by the overwhelming numerical predominance of the English root + English affix combination in most of these cases. For example *branding* and *rebranding* are used for more than 80 times in *Capital* 2005, while *rebrandare* for merely 10, *marketing* has over 850 occurrences, while *marketare* appears once, *retailer* has over 120 occurrences, while *retailist* is used once. These findings lend support to those interpretations positing a direct connection between blended derivatives and their source language models (E. Haugen 1956, M. Avram 1997).

---

[12] E. Haugen, 1956, p. 386.

However, the English stems in many of the loanblends discussed above show a very high frequency of occurrence, being among the top 100 most frequent types in the studied corpus: *brand* alone is used for 420 times*, retail* has a token frequency of 350, while *target* occurs in 32 instances, to give just a few examples. Moreover, sometimes the Romanian suffix replaces no foreign morpheme, entering borrowings which lack derived models in English: *PR-isti (PR representatives), softist (software engineer), mixaj (mix), sponsoriza (sponsor*), *masterand (master's student) masterat (master's programme)*. This situation seems to offer support to those theoretical positions (D. Winford, 2003) which explain hybrid loanwords as resulting from recipient language derivational processes applied to previously imported words, rather than being modeled on equivalent derivatives in the source language. From this perspective, it would be tempting to explain *retailist,* for example, not as an adapted form of *retailer*, but as the Romanian-internal derivative *retail + ist*, on the model of *detailist, angrosist, profesionist, etc*. However, we agree with Haugen that "*we cannot check the mental processes of the speakers*", and that in situation where "*several factors have operated, we have no good way of saying which one was the most important*"[13].

Hybrid compounds or blended compounds in Haugen's taxonomy are those words consisting of both source and recipient language stems. Substitution in blended compounds, Haugen shows, requires the speaker's ability to analyze the model he is imitating. For example PaG (Pennsylvanian German) adopted AmE *plum pie* as [blaumapai], because the German speaker, being aware of the compositional nature of this word, was able to 'break' it into component parts and import the English *pie* but substitute the native *blaume* for 'plum'[14]. Clyne[15] presents a number of similar examples from the German spoken in Australia, e.g. *Gumbaum-* gumtree, *Redbrickhaus-* red brick house, *Grungrocer-* greengrocer, *Frontgarten-* front garden, *Lunchzeit-* lunchtime. Such combinations seem to be supported by a formal similarity existing between the foreign word and the native one which is substituted, although Weinreich[16] also reports on blended compounds clearly diverging from the source language models on which they were formed: PaG *Esixjug- vinegar jug* and *fleischpie-* meat pie, Spanish *pelota de fly -* fly ball.

The Romanian corpus studied in this paper contains examples both of phonetically motivated compound blends (*crash-teste-* crash-tests*, business-planuri-* business-plans, *masterfranciză-* masterfranchise, *schipass-* skipass, *spray-*

---

[13] *Ibidem*, p. 441.
[14] *Ibidem*, p. 390.
[15] Michael CLYNE, 1967, *Transference and triggering: Observations on the language assimilation on postwar German-speaking migrants in Australia*, The Hague: Martinus Nijhof, p. 35.
[16] Uriel WEINREICH, 1968, *Languages in Contact: Findings and Problems*, The Hague, Paris: Mouton, p. 52.

101

*pompa-* spray-pump, *team-lider-* team-leader, *hair-stilist-* hair-stylist, *brand-umbrela-* umbrella brand) and of blended compounds in which one element is formally unrelated to the model *(brand-mamă* – parent-brand, *vacante-bonusuri-*bonus holidays). The classification of possible blended compounds such as *audiobook, videochat, videoplayer, videostreaming, videoclip* is complicated by the use of the bilingual homophones *audio* and *video,* which in the absence of pronunciation cues are difficult to mark as belonging to either English or Romanian. However, Weinreich[17] believes that when one element of a hybrid compound is affected by homonymy, this will be rendered by the homophonous native word with a slightly extended meaning, while the other one will be transferred as such, or 'imported' to use Haugen's terminology. In the light of this proposal, we have chosen to treat the above examples as blended rather than as English compounds.

Turning now to the various structural patterns these compounds use, noun + noun combinations are by far the most commonly employed ones, followed by adjective + noun and verb + noun combinations. Within the first group, the most frequent situation occurs when a nominal English modifier combines with a Romanian head. These are marginal loanblends in Haugen's classification, e.g. *babyschi, focus-grup, masterfranciza, media-plan, vacante-bonusuri, crash-teste, hair-stilist.* Nuclear compounds, i.e. compounds in which the head is borrowed while the modifier is Romanian, include *brand-umbrela, brand-mama, audiobook, videochat, videostreaming, videoclip, videoplayer, schipass.* However, as seen from the examples above, these compounds have their semantic and grammatical head on the rightmost nominal element, while Romanian usually has it on the leftmost word in the compound.

### 3.3. Loanshifts

Loanshifts are words which show morphemic substitution without importation, or words in which the meaning is imported without the foreign form. They can result from the extension of a meaning in the recipient language so as to correspond to that of a word in the source language (*semantic loans* or *semantic extensions*), or from the importation of a morpheme arrangement from this language (*loan translations* or *calques*).

Semantic loans are most often motivated by "*both phonetic and semantic resemblance between foreign and native terms*"[18]. For example, Clyne[19] shows that German/English bilinguals in Australia have taken the German word *magasin* meaning 'storeroom', and have extended its meaning to that of the English word *magazine*. Similarly, Weinreich[20] describes the case of the American Italian word

---

[17] *Ibidem.*
[18] E. Haugen, 1950, p. 92.
[19] Clyne, 1967, p. 61.
[20] Weinreich, 1968, p. 49.

*fattoria*, which originally meant 'a farm', but which later came to mean *factory*. Examples of semantic loans reported in the Romanian literature (M. Avram, 1997; A. Stoichitoiu-Ichim, 2001; Constantinescu, Popovici and Stefanescu 2004), or found in our corpus are: *domestic* as 'tame, living near human habitations', which received the new meaning of 'national, internal'; *a aplica* as 'to put into operation or effect', which has also come to mean 'to make a request or appeal'; *a realiza* as 'to accomplish', which has received the additional meaning 'to be aware of'; *maturitate* as "the quality of being mature, full development", which has also come to mean "termination of the period that an obligation has to run", and many others.

The second subclass of loanshifts in Haugen's classification (i.e. creations, also called 'loan translations' or 'calques') are most often found at the level of compounds. For example, Romaine (1995, p. 57) shows that the English *skyscraper* was borrowed in different languages as a rearrangement of native morphemes: *gratteciel* in French, *rascalielos* in Spanish, *Wolkenkratzer* in German, etc. In recent years Romanian has calqued a number of English expressions such as 'first lady' *prima doamnă*, 'no man's land' *ţara nimănui*, 'number one' *numărul unu,* 'second hand' *la mâna a doua*[21], 'brainwashing' *spălarea creierelor*, 'human rights' *drepturile omului*, 'flying saucer' *farfurie zburătoare*[22].

### 4. Conclusions and outlook

The analysis of borrowing from English into present-day Romanian has revealed the richness and complexity of this process. Thus, from assimilated to unassimilated loans, from pure to blended or merely semantic transfers, the language of the studied corpus provides examples from all the main classes of borrowings in Haugen's classical taxonomy. However, the relationship between these classes deserves further study. Although there is some evidence which shows that sometimes already adapted borrowings are reintroduced into the language in an unadapted form, the relationship between loanwords and loanshifts is less clear. Recent studies on the topic (Z. Manolescu, 1999) have shown that the number of English loanwords in present-day Romanian is on the increase. In this context, it would be interesting to see whether this increase is paralleled by a rise in the number of semantic loans, or takes place at the expense of this class of borrowings.

**Bibliography**
AVRAM, Mioara, 1997, *Anglicismele în limba română actuala*, Conferinţă prezentată la Academia Română. Bucureşti: Editura Academiei Române.
CIOBANU, Georgeta, 1996, *Anglicismele în limba română*, Timişoara: Amphora.

---

[21] M. Avram, 1997, p. 27.
[22] I.Constantinescu, V. Popovici, A.Ştefănescu, 2004, „Romanian", in GÖRLACH, Manfred (ed.), pp. 189-190

103

CLYNE, Michael, 1967, *Transference and triggering: Observations on the language assimilation on postwar German-speaking migrants in Australia*, The Hague: Martinus Nijhof.

CONSTANTINESCU, I.;POPOVICI, V.; ȘTEFĂNESCU A., 2004, „Romanian”, in GÖRLACH, Manfred (ed.), pp. 168-194.

GÖRLACH, Manfred (ed.), 2004, *English in Europe*, New York: Oxford University Press.

HAUGEN, Einar, 1950, "The Analysis of Linguistic Borrowing", in: *Language*, 26 (2), pp. 211-231

HAUGEN, Einar, 1956, *Bilingualism in Americas: A Bibliography and Research Guide*, University of Alabama Press.

MANOLESCU, Zoia, 1999, *The English Element in Contemporary Romanian*, București: Conspress.

MYERS SCOTTON, Carol, 1993, *Duelling Languages: Grammatical Structure in Codeswitching*, New York: Oxford University Press.

ONYSKO, Alexander, 2007, *Anglicisms in German: Borrowing, Lexical Productivity, and Written Codeswitching*, Berlin, New York: Walter de Gruyter.

POPLACK, Shana, 1980, „Sometimes I'll start a sentence in Spanish y termino en español: toward a typology of code-switching", in: *Linguistics* 18, pp. 561-618.

ROMAINE, Suzanne, 1995, *Bilingualism*, Oxford: Blackwell Publishers.

STOICHIȚOIU-ICHIM, Adriana, 2001, *Vocabularul limbii române actuale: Dinamică, influențe, creativitate*, București: Editura All educational.

STOICHIȚOIU-ICHIM, Adriana, 2006, *Aspecte ale influenței engleze în româna actuală*, București: Editura Universității din București.

TREFFERS-DALLER, Jeanine, 2000, "Borrowing", in *Verschueren*, Jef, Jan-Ola Östman, Jan Blommaert, Chris Bulcaen (eds.).

VERSCHUEREN, J.; ÖSTMAN, J,; Blommaert, J.; Bulcaen, C. (eds.), 2000, *Handbook of pragmatics*, Amsterdam, Philadelphia: John Benjamins.

WEINREICH, Uriel, 1968, *Languages in Contact: Findings and Problems*, The Hague, Paris: Mouton.

WINFORD, Donald, 2003, *An Introduction to Contact Linguistics*, Blackwell Publishing.

Colecția CD Capital 2005, București: Ringier România.