

Negative affixation in native and non-native English: An onomasiological corpus-based approach

1. Background of the study

In the field of second language acquisition, learners and teachers alike have always been in search of the magic potion that would make it possible to master a second/foreign language and use it in a way that is, if not identical to, at least sufficiently similar to native speakers' use so as to be able to join their linguistic community. In this regard, the advent of corpora representing both native and non-native language has helped linguists get a better understanding of what distinguishes the two varieties and hence how learners can improve their proficiency in the language (and teachers can help them do so). Linguistic research has thus underlined several aspects that could benefit from more attention in language teaching, be it in the domain of syntax (e.g. gender agreement in L2 French in Bartning 2000), phraseology (e.g. verb-noun collocations in Nesselhauf 2005) or pragmatics (e.g. German modal particles in Belz & Vyatkina 2008). One aspect that has been relatively underresearched up until now is derivational morphology (for an exception, see Callies 2015). Yet, it cannot be denied that morphology, just like syntax or vocabulary, can contribute to the “native soundingness” or “non-native soundingness” of learner production. In an attempt to move this line of research forward, the present paper adopts an onomasiological approach by focusing on the expression of negation in native and non-native English through the use of affixes (see Cartoni & Lefer 2011 for a similar approach in a cross-linguistic perspective). It relies on corpus data that represent the authentic written production of students from different mother tongue backgrounds, including native British students. At the same time, the study illustrates the recent trend that consists in comparing EFL and ESL varieties (cf. Nesselhauf 2009, Gilquin 2015), since it explores the derivational morphology of students for whom English is a foreign language (mainland Chinese and Belgian French-speaking learners) as well as those for whom English is an institutionalized second-language variety (Hong Kong and Singapore students).

2. Data and methodology

All the corpus data correspond to student academic writing. The native data (ENL) and the ESL data come from the International Corpus of English (ICE; Greenbaum 1996), while the EFL data come from the International Corpus of Learner English (ICLE; Granger et al. 2009). Table 1 presents the five different (sub)corpora and the number of words each of them contains.

Table 1. Corpora and word counts

	Corpus	Word count
ENL	ICE-Great Britain (GB)	43,379 words
ESL	ICE-Hong Kong (HK)	49,420 words
	ICE-Singapore (SIN)	46,435 words
EFL	ICLE-Chinese (CH)	95,073 words
	ICLE-French (FR)	190,544 words

Six affixes were selected for investigation because they can all be used to express negation in English: the prefixes *de-*, *dis-* (and its allomorphs *di-* and *dif-*), *in-* (and its allomorphs *il-*, *im-* and *ir-*), *non-*, *un-*, as well as the suffix *-less*. The combinations of letters corresponding to these affixes were extracted automatically from the corpora by means of WordSmith Tools (Scott 2008) and the output was then manually

examined to remove all cases where the search string did not function as a negative affix (semantically opaque cases were disregarded too). At this stage the prefix *a-* was not included in the study because of its purported lack of productivity (see Bauer 1983). The absolute frequency, relative frequency, percentage and rank of each affix in the various corpora were computed, and differences between the native and non-native results were tested for significance by means of a log-likelihood ratio statistic. In addition to the six affixes, the negative particle *not* was searched for as a way of evaluating the degree to which negation is expressed in a non-derivational fashion.

3. Results

3.1. Affixation in ENL, ESL and EFL

As a rule, it appears that non-native writers underuse negative affixation as compared to native writers, in a way that is statistically significant ($p < 0.005$) – cf. Figure 1. This underuse is even stronger in EFL than in ESL, which one might explain by the degree of exposure to the target language: EFL learners mainly get access to English through instruction, in the environment of school or university. In ESL countries, on the other hand, English serves as an official or semi-official language and hence is also used in certain everyday contexts (for example in the media or in the administration), which results in increased exposure to naturally-occurring language for ESL students and a better chance of encountering (negative) affixes and gradually assimilating them.

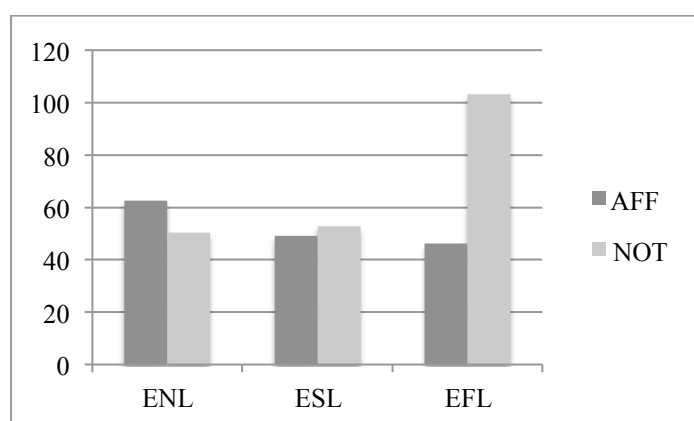


Figure 1. Relative frequency (per 10,000 words) of negative affixation (AFF) and negative particle (NOT) in ENL, ESL and EFL corpora

Interestingly, while non-native writers use affixation less often than native writers, they use the negative particle *not* more often (this is particularly striking among EFL learners). What this suggests is that non-native writers might compensate for the underuse of negative affixes by using the negative particle instead. This is illustrated by (1), where the learner has used the phrase *not capable*. While this is a perfectly acceptable way of expressing negation, a search on the large reference corpus that is the British National Corpus reveals that *incapable* is more likely to occur than *not capable* (818 occurrences of *incapable* vs 141 occurrences of *not capable*).

- (1) Africa with its specific climatic conditions is **not capable** of challenging European agricultural productivity. (ICLE-FR-ULB-0008.2)

3.2. Affixes and learner populations

The above tendencies, namely an underuse of negative affixes and an overuse of the negative particle, are valid for all the learner populations represented in the corpora (although the difference is not always statistically significant). When it comes to the

distribution of the individual affixes, on the other hand, there is some variation between the learner populations. As Table 2 shows, *un-* and *in-* are the two most common affixes, but while *un-* is the top affix in native English as well as Hong Kong and mainland Chinese English, in Singapore and (Belgian) French English it is *in-* that comes first. Interestingly, if we look at the proportion of these two affixes, we can notice two clusters: one with GB, SIN and HK, in which the top affix represents over 40% and the top two affixes make up close to 80% of the total, and one with FR and CH, in which there is a slightly more uniform distribution of the affixes, with a less frequent top affix (36%) and the thirdly ranked affix (*dis-*) representing a higher percentage than in any of the other corpora. This difference between the EFL and ESL varieties might, again, be related to the students' degree of exposure to the target language: the more exposure they receive, the more likely they seem to reproduce the types of frequency distribution that are found in native English.

Table 2. Percentage of the negative affixes in the five corpora

GB	SIN	HK	FR	CH
1. <i>un-</i> (44%)	1. <i>in-</i> (41%)	1. <i>un-</i> (44%)	1. <i>in-</i> (36%)	1. <i>un-</i> (36%)
2. <i>in-</i> (33%)	2. <i>un-</i> (33%)	2. <i>in-</i> (36%)	2. <i>un-</i> (32%)	2. <i>in-</i> (24%)
3. <i>dis-</i> (13%)	3. <i>dis-</i> (12%)	3. <i>dis-</i> (9%)	3. <i>dis-</i> (17%)	3. <i>dis-</i> (23%)
4. <i>non-</i> (4%)	4. <i>non-</i> (7%)	4. <i>non-</i> (6%)	4. <i>-less</i> (8%)	4. <i>-less</i> (14%)
5. <i>-less</i> (3%)	5. <i>-less</i> (5%)	5. <i>de-</i> (3%)	5. <i>de-</i> (4%)	5. <i>non-</i> (2%)
6. <i>de-</i> (3%)	6. <i>de-</i> (2%)	6. <i>-less</i> (2%)	6. <i>non-</i> (3%)	6. <i>de-</i> (1%)

3.3. Transfer from the mother tongue

The inter-varietal differences mentioned above may trigger the question of whether such differences could be due to the influence of the students' mother tongue background. Focusing on ICLE-FR, for which we know from the corpus metadata that all the participants had French as their mother tongue, we could argue that the preference for *in-* above all other negative affixes and the higher proportion of *dis-* and (to a much lesser extent) *de-* as compared to native English could be due to the influence of French, since these four affixes are of Romance origin. Yet, because some of these features are shared by the other populations (whose L1 background is uncertain for lack of corpus metadata, but who are likely to be native speakers of some variety of Chinese), it is difficult to demonstrate the existence of L1 influence on the basis of frequency alone. A close examination of the corpus data themselves, however, makes it clear that L1 transfer is a plausible explanation for some of the non-standard forms produced by the French-speaking students. Consider example (2), where the use of the word *deshumanised*, instead of the correct form *dehumanised*, is probably the result of transfer from the French equivalent *déshumanisé*.

- (2) They are living in a **deshumanised** world, where only productivity matters, where human relations and family life have become less important than money. (ICLE-FR-UCL-0091.3)

Quite surprisingly, the ICLE-FR data also contain non-standard (or dispreferred) forms that go against the principles of French morphology. Thus, the use of *unrational* in example (3) is unexpected, since the correct form, *irrational*, has a direct counterpart in French, *irrationnel*. Similarly, it is surprising that in (4) a French-speaking learner should have produced the form *unconceivable* (listed in dictionaries as a possible word of the English language, but not occurring a single time in the whole British National Corpus), rather than *inconceivable*, given the

French equivalent *inconcevable*. In such cases, other mechanisms than L1 transfer are clearly at work, like – possibly – a process of overgeneralization of the prefix *un-* to express negation (see Callies 2015: 140), also visible in example (5) from ICLE-CH.

(3) The result of this **unrational** way of thinking is that the world appears as a chaos. (ICLE-FR-UCL-0048.2)

(4) So, such a social aid is **unconceivable**. (ICLE-FR-UCL-0086.3)

(5) Human life is **unreplaceable**, we should not carry out the capital punishment because of 'fairness' to those were killed. (ICLE-CN-UK-0105.1)

4. Concluding remarks

On the basis of the first results outlined above, it appears that non-native writers' use of negative affixation can differ from that of native writers, both from a quantitative and qualitative point of view. Among the possible explanations for some of these differences are the influence of the mother tongue and the activation of general cognitive processes such as overgeneralization. The study has also confirmed the relevance of comparing EFL and ESL varieties, since they are sufficiently similar to allow for comparison and at the same time present interesting differences which can partly be related to the contexts of acquisition of these two varieties.

This research will be expanded to include aspects like type-token ratio, presence of hapax legomena or degree of innovation, as well as other language varieties, including spoken English. Not only has such an investigation the potential to advance our understanding of interlanguage morphology, but it might help uncover important cognitive mechanisms underlying the acquisition of a foreign/second language.

References

- Bartning, I. 2000. 'Gender agreement in L2 French – preadvanced vs advanced learners'. *Studia Linguistica* 54(2): 225–37.
- Bauer, L. 1983. *English Word-Formation*. Cambridge: Cambridge University Press.
- Belz, J. A. and N. Vyatkina. 2008. 'The pedagogical mediation of a developmental learner corpus for classroom-based language instruction'. *Language Learning and Technology* 12(3): 33–52.
- Callies, M. 2015. 'Effects of cross-linguistic influence in word formation. A comparative learner corpus study of advanced interlanguage production'. In H. Peukert (ed.) *Transfer Effects in Multilingual Language Development*. Amsterdam: Benjamins, 127–43.
- Cartoni, B. and M.-A. Lefer. 2011. 'Negation and lexical morphology across languages: Insights from a trilingual translation corpus'. *Poznan Studies in Contemporary Linguistics* 47(4): 795–843.
- Gilquin, G. 2015. 'At the interface of contact linguistics and second language acquisition research: New Englishes and Learner Englishes compared'. *English World-Wide* 36(1): 90–123.
- Granger, S., E. Dagneaux, F. Meunier and M. Paquot. 2009. *The International Corpus of Learner English, Version 2. Handbook and CD-ROM*. Louvain-la-Neuve: Presses universitaires de Louvain.
- Greenbaum, S. (ed.). 1996. *Comparing English Worldwide: The International Corpus of English*. Oxford: Clarendon Press.
- Nesselhauf, N. 2005. *Collocations in a Learner Corpus*. Amsterdam: Benjamins.
- Nesselhauf, N. 2009. 'Co-selection phenomena across New Englishes'. *English World-Wide* 30(1): 1–26.
- Scott, M. 2008. *WordSmith Tools version 5*. Liverpool: Lexical Analysis Software.