

# Sorting out translation universals from specific source-language interference

The case of phrasal verbs in translated English

Rudy Loock and Bert Cappelle

University of Lille 3 & National Center for Scientific Research (CNRS), France

# Phrasal verbs are typical of English

“Throughout the history of English phrasal verbs have always had a place in the ‘**common core**’ of the language.”

(Thim 2012: 244)

# Phrasal verbs tend to be colloquial

“Phrasal verbs are mainly **colloquial**. They are used casually, in everyday speech, or in order to express vivid, emotional and frequently slangy points, to conjure up special metaphoric relationships and jokes [...]”

(McArthur & Atkins. 1974. *Dictionary of English Phrasal Verbs and their Idioms*, cited in Chuquet & Paillard 1987: 199)

“redundant”, “improper”, “low”, “barbarous”, “vulgar”, “less elegant”, ...

(Samual Johnson. 1799. *A Dictionary of the English Language*, cited in Wild 2010: 207)

F\*\*\* off!

# An illustration

The previous slide, you will agree, contains an expression which is

- typical of English lexis and grammar
- very rude, not part of what is the standard

# Two opposite predictions based on TUs

## 1. Normalization / exaggeration of target language norms

“to make a text sound English, make sure to use lots of phrasal verbs”

→ MORE phrasal verbs in translated English than in original English

## 2. Standardization

“better be safe and not use an expression that sounds too colloquial”

→ FEWER phrasal verbs in translated English than in original English

# Problem

These TU-based predictions only focus on the target language

They ignore this important question:

How readily does a phrasal verb present itself as a translation **of what's in the source text?**

→ Contrastive issues should be taken into account

# A previous study

Cappelle (2012):

translated English contains fewer manner-of-motion verbs when source language is French than when source language is German

Reason:

- English & German: satellite-framed (direction in a particle / PP)

A UFO whizzed by.

Ein UFO sauste vorbei.

- French: verb-framed (direction in the verb)

Un OVNI passa (à grande vitesse).

# Research question

If we focus on particles rather than on motion verbs, do we see a similar source-language effect?

# Method

- Two items selected: *up* and *out*
  - these words are likely to be used as a particle
  - they're the most frequent particles

(Gardner & Davies 2007)

- Using large corpora
  - Translational English Corpus (TEC): you can select different source languages
  - British National Corpus (BNC): a 'reference corpus' (100 million words)
- Rather crude method
  - no manual sifting of results
  - not even manual precision and recall tests for the queries (in principle still possible)

# Data

## TEC<Rom

- English fiction since 1990 translated from
  - French
  - (European) Spanish
  - Latin American Spanish
  - (European) Portuguese
  - Brazilian Portuguese
  - Catalan
- 1,258,951 words

## TEC<Ger

- English fiction since 1990 translated from
  - Danish
  - Dutch
  - German
  - Icelandic
  - Norwegian
  - Swedish
- 1,008,028 words

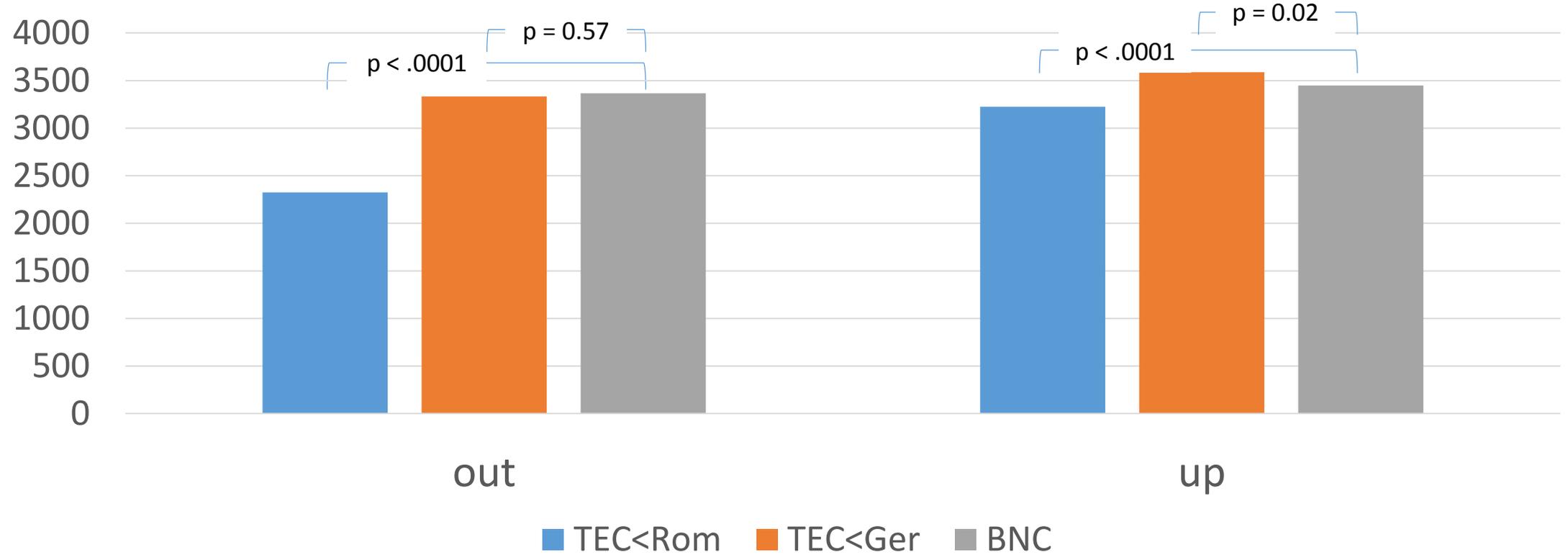
# Data (cntd.)

## BNC

- Subcorpus of fiction selected
- 15,909,312 words

# Results

Corpus occurrences (per million words)



# Results in words

- Both *up* and *out* are highly significantly underrepresented in the corpus of English translated from Romance languages
- *Out* has almost the exact same frequency in the corpus of English translated from Germanic languages as in non-translated English
- *Up* is even slightly more frequent in the corpus of English translated from Germanic languages than in non-translated English, but only at a low threshold of significance

# Discussion

- The frequency of particles in translated English is not across-the-board over-represented (due to e.g. ‘normalization’) or under-represented (due to e.g. ‘standardization’)
- There is a clear, highly significant difference in the frequency of these particles depending on the kind of source language (Germanic vs. Romance)
- This result pleads in favour of source-language interference as more relevant than the above-mentioned Translation Universal candidates

# Discussion (cntd.)

- If there IS a Translation “Universal” at play here, it’s the Unique Items hypothesis, which is formulated in terms of pairs of **specific** languages  
(Tirkkonen-Condit 2002; Kolehmainen 2013)
- Language interference ties in with the findings of some of our previous studies  
(Cappelle 2012; Cappelle and Loock 2013; Loock, De Sutter & Plevoets 2013)

# Further question

What about the translations from Romance that do make use of a particle?

→ Hypothesis:

While some of them may also have something particle-like (e.g. a prefix), most of them will be simplex (explaining the underrepresentation effect)

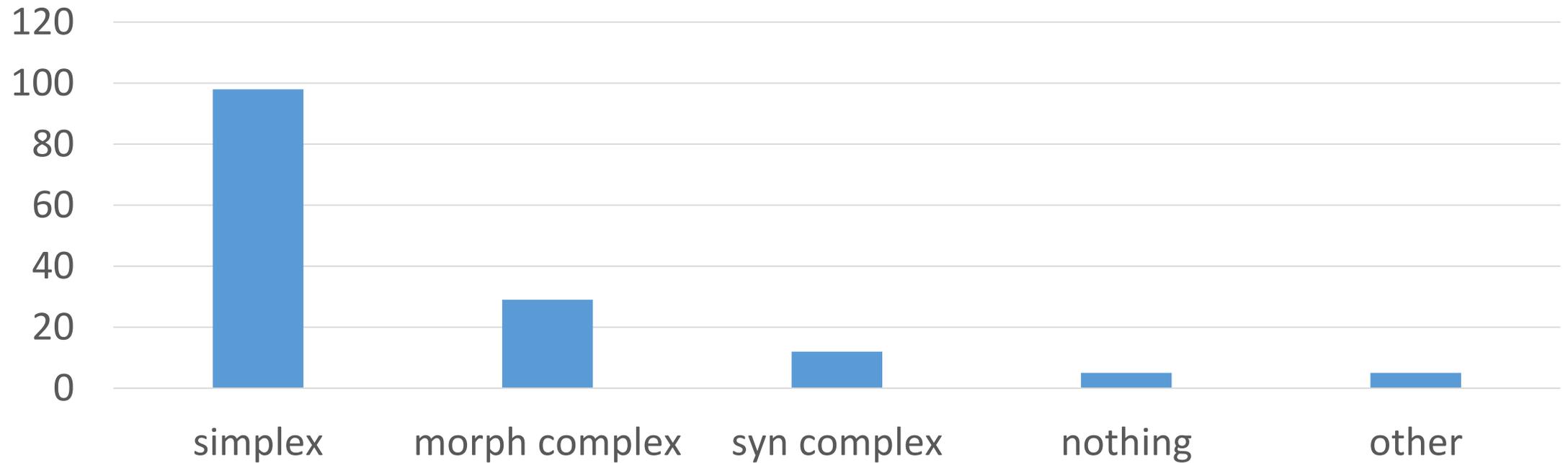
- TEC doesn't show the source texts
- So, we will need a different, more time-consuming methodology: comparing source and target texts

# More qualitative follow-up study: method

- *Le Petit Prince* (A. de Saint-Exupéry, 1943) and its English translation
- 14,952 and 17,066 words, respectively (+ 14%!)
- Align the texts
- Look up all the particles in the translation
- Look up the corresponding expressions in the source text
- Code these for structure:
  - simplex (*abandonner* → *give up*)
  - syntactically complex (*dire encore* → *go on*)
  - morphologically complex (e.g. *re-venir* → *come back*)
  - (altogether different formulation ('other') / nothing in source)

# Results

Structure of source expression in *Le Petit Prince* corresponding to a verb-particle combination



# Discussion

Most phrasal verbs in *The Little Prince* have a simplex source expression

- This pleads for the Unique Items hypothesis
- This also leads to the hypothesis:

most phrasal verbs in translations from Germanic languages have **complex** source expressions

→ needs to be confirmed

- Further hypothesis:

simplex items in Romance languages are more likely to be translated by a simplex item in English than complex items are

→ needs to be confirmed

# Discussion (cntd.)

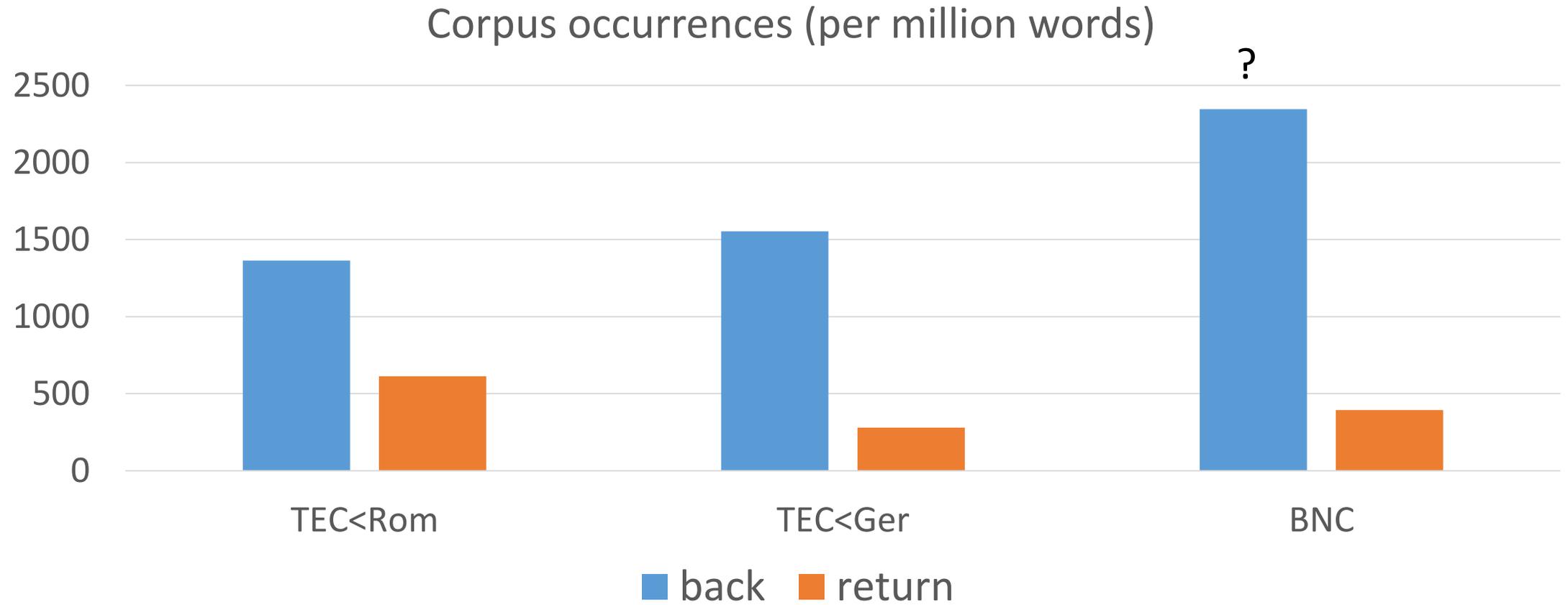
Most morphologically complex expressions in the source text have *back* rather than any other particle as a translation (19/29 cases)

Reason: *re-* is a common prefix in French

→ Hypothesis:

maybe *back* is not underrepresented in TEC<Rom

# But it is...



# Conclusions on methods

- No advanced statistical technique was used here: just Chi-square
  - If anything, the advanced nature of our study lies in the use of rather large text samples
    - ranging from over a million to almost sixteen million words
    - significant results are easily obtained
    - it's harder to find something that is NOT significant (but we found that, too!)
  - We plead for combining large-scale 'automatic' quantitative research with manually-coded, more qualitative research
- each method drives hypotheses to be tested with the other one

Thank you!

Correspondence: [Rudy.Loock@univ-lille3.fr](mailto:Rudy.Loock@univ-lille3.fr) / [Bert.Cappelle@univ-lille3.fr](mailto:Bert.Cappelle@univ-lille3.fr)

Check out CorTEx website: <http://stl.recherche.univ-lille3.fr/CorTEx/>

# References (other than those in the abstract)

- Cappelle, B. and R. Loock. 2013. Is there interference of usage constraints? A frequency study of existential *there is* and its French equivalent *il y a* in translated vs. non-translated texts. *Target* 25:2, 252-275.
- Chuquet, H. and M. Paillard. 1987. *Approche linguistique des problèmes de traduction anglais-français*. Paris: Ophrys.
- Kolehmainen, L. 2013. Die Unikat-Hypothese der Translation: Etwas Altes, etwas Neues und etwas Geliehenes. *Trans-kom* 6:2, 92-114.
- Loock, R., G. De Sutter and K. Plevoets. 2013. Teasing apart Translation Universals and Source-Language Interference: a case study on derived adverbs in English and French. Talk presented at ICLC 7 - UCCTS 3, Ghent, 11-13 July 2013
- Tirkkonen-Condit, S. 2002. Translationese – a myth or an empirical fact? A Study into the linguistic identifiability of translated language. *Target* 14(2): 207-220.
- Wild, C. 2010. Attitudes towards English usage in the late modern period: the case of phrasal verbs. Unpublished PhD thesis, University of Glasgow.