

# Translation Corpus-Informed Research: A Swedish-Croatian Example



Goranka Antunović

**Abstract** Corpora including original texts in one language and their translations in another (translation corpora, also called parallel corpora) have proved a useful basis for contrastive studies as well as for research in the field of Translation Studies and their potential as well as limitations have already been discussed in the literature (Aijmer, 2007; Granger, 2003, etc.). They help establish relationships between units of the source and the target language, enable better understanding of the meaning and the usage of those units but they can also shed light on various translation-related phenomena, those specific for the language pair involved as well as more general. This cross-linguistic study seeks to establish Croatian translation equivalents of the Swedish indefinite pronoun *man* by analysing a >410,000-word corpus of authentic Swedish texts and their Croatian translations. These are compared to the standard lexicographic equivalents of the pronoun in an attempt to contribute to a more complete presentation of the Croatian correspondents of *man*. Remarks are made regarding the potential contribution of the findings to a better understanding of the usage of the Swedish pronoun as well as to the study of translation equivalence and translators' preferences.

**Keywords** Translation corpora · Contrastive analysis · Translation studies  
Swedish · Croatian · Indefinite pronoun *man*

## 1 Background

Text corpora of various kinds, among them those composed of original texts in one language and their translations in another, play a significant role in Contrastive Linguistics (CL) as well as in Translation Studies (TS) and have “the potential to bring the two fields even closer together” (Granger, 2003, p. 25). CL has traditionally seen

---

G. Antunović (✉)

Scandinavian Studies Section, English Department, Faculty of Humanities  
and Social Sciences, University of Zagreb, I. Lučića 3, 10000 Zagreb, Croatia  
e-mail: [gantunovic@ffzg.hr](mailto:gantunovic@ffzg.hr)

© Springer Nature Switzerland AG 2019

B. Lewandowska-Tomaszczyk (ed.), *Contacts & Contrasts in Educational Contexts  
and Translation*, Second Language Learning and Teaching,  
[https://doi.org/10.1007/978-3-030-04978-2\\_10](https://doi.org/10.1007/978-3-030-04978-2_10)

corpora as essential (cf. Filipović, 1973), and its revival over the last three decades has often been attributed to the development of various electronic tools, in particular computer corpora as one of the two key factors (the other one being globalization; cf. Gómez-González & Doval-Suárez, 2005; Granger, 2003; Ramón García, 2002). The new context has affected the discipline in various ways, including, as Ramón García (2002) notes, “a shift in the focus of CL from linguistic systems in a more abstract sense to more specific issues in language use” (p. 395). The great potential of corpus-based research in the field of TS has also been identified by many researchers, who have discussed opportunities as well as limitations (among early examples are Altenberg & Aijmer, 2000; Baker, 1993, 1995; contributors to *META 43(4)*, 1998). Ramón García (2002) thus states that “[e]lectronic corpora have been used extensively in TS since the beginning of the 1990s and their importance is such that they have meant a turning point in the discipline” (p. 401). Corpora have been seen as useful not only in relation to machine translation and terminology but also in studies exploring issues such as translation universals, language of translation, translator style, translation ideology, translation strategies and methods, equivalence, cognitive processes in translation, etc. (cf. Altenberg & Aijmer, 2000, p. 27; Lan, 2014; Lewandowska-Tomaszczyk, 2012b; Mikhailov & Cooper, 2016, pp. 16, 184–192).

The two disciplines do not only share the need to make use of corpora but they are closely intertwined in a more general sense, as was noted already in the decades preceding the appearance of electronic corpora (cf. Ramón García, 2002, p. 397), as well as later on. Hoey and Houghton (2001) describe the bidirectional relationship between Contrastive Analysis (CA) and translation by pointing out that “(...) the translation of specific pieces of text may provide the data for CA (...). CA may provide explanations of difficulties encountered in translation” (p. 49). Just as they strongly state that “Translation as a source of data for CA is strictly unavoidable”. (ibid.), Ramón García (2002) concludes that “(...) any type of approach to translation from a descriptive corpus-based perspective must take into account some kind of contrastive aspect. (...) CL is thus a basic ingredient of TS” (p. 403).

The present study wishes to identify some insights in the realm of either CL or TS that can be gained from focusing on a particular linguistic unit, namely the Swedish pronoun *man*, and its Croatian translation equivalents in a translation corpus. The possibility to use one and the same corpus as a basis for conclusions within both of the two closely connected disciplines seems particularly welcome in cases such as this, where lesser-taught languages and lesser-studied language pairs are involved.

## 2 Translation Corpora

The corpus used in the present study consists of original texts in Swedish and their translations into Croatian and it is termed here a ‘translation corpus’, in line with e.g., Altenberg and Aijmer (2000, p. 16). In the literature such corpora are referred to as either ‘translation corpora’ or ‘parallel corpora’ (or both, as in e.g., Lewandowska-Tomaszczyk, 2012a, p. 5; Ramón García, 2002, p. 401). Granger (2003) has suggested

that the choice of the name may depend on the affiliation of the researcher ('parallel corpora' in the TS framework, 'translation corpora' in CL) but that such usage "is not entirely consistent either" (pp. 19–20).

Translation corpora offer good grounds for cross-linguistic research for a number of reasons. They provide a large number of pairs of related units in different languages, which enables comprehensive and nuanced conclusions about cross-linguistic correspondences. The fact that for one source-language unit a relationship is typically established with a number of different target-language renderings does not only "make it possible to discover cross-linguistic variants" (Altenberg & Granger, 2002, p. 9) but it also sheds light on different meanings and different functions of the units thus enabling their better understanding within their respective languages. Even if the exact nature of the relationship established cannot be unequivocally defined—which in itself is a point of great interest for researchers—the fact that two units are seen as establishing translation equivalence in numerous occurrences justifies conclusions about their correspondence and invites its careful interpretation. For TS researchers corpora are a rich source of authentic translation solutions which enable them to "captur[e] the distinctive features of the translation process and product" (Granger, 2003, p. 22). Another advantage of translation corpora, no less significant for being obvious, is that they make research more objective by providing "access to a large number of interpretations besides the linguist's own introspective judgement" (Aijmer, 2007, p. 33).

While they undoubtedly present a most useful tool, translation corpora have their limitations, and conclusions based on their evidence must be made with caution. Granger (2003), when discussing limitations with regard to corpus-based approach in general, focuses on two such circumstances, namely the limited availability of corpora (still true today of large translation corpora for many language pairs, including Swedish-Croatian) and "the corpus predilection for form-based research" (pp. 22–23). While these two may influence the decision on the kind of study to be conducted, there are other characteristics of translation corpora that may influence the quality of the study and the value of its findings. They stem from the very nature of translation process and translation product, and primarily have to do with the language of translations and the elusive nature of 'translation equivalence'. When pairs of linguistic units are analysed one must bear in mind that different language systems will possess "what can be called equivalents not of an *identity* but rather of an *approximation* type" (Lewandowska-Tomaszczyk, 2012b, p. 4). Researchers using translation corpora must also recognize that translations tend to be influenced by the source text (so much so that a 'law of interference' has been proposed by Toury, 1995, p. 275), and that this results in phenomena that may be very difficult to spot (e.g., lexical frequencies different from those in original texts in the same language, as has been observed by Gellerstam, 1986 and others). Translation scholars have identified other common features of the language of translated texts, not source-inspired, e.g., selection of linguistic units in line with what Toury (1995) has termed 'the law of growing standardization' (p. 267). Awareness of these, as well as of the intricate nature of translation equivalence (including the very basic distinction

between translation equivalence and contrastive correspondence, cf. Ivir, 1997) may prove essential in securing sound conclusions based on evidence from translation corpora.

### 3 Swedish Pronoun *Man* and Its Croatian Equivalents

#### 3.1 *Man*

Descriptions of the pronoun *man* in grammars of the Swedish language show that its usage is by no means simple. It is often classified as an indefinite pronoun (Bolander, 2005; Holmes & Hinchliffe, 2003; Ljung & Ohlander, 1971/1982) or, more specifically, as a quantitative indefinite pronoun (Josefsson, 2009/2011, who remarks (p. 82) that the Swedish Academy's Grammar also classifies *man* as quantitative). Hultman (2003/2008) describes it as a generalizing quantitative pronoun but also points to its anaphoric use "when the speaker/writer cannot or doesn't want to make his/her reference precise" (p. 122).<sup>1</sup> In a similar way the generalizing function is implied in other grammars (e.g., Bolander, 2005, p. 128), often explicitly or implicitly connected with remarks about the pronoun, in itself singular, appearing with plural complements in colloquial Swedish (cf. Bolander, 2005, pp. 127–128; Holmes & Hinchliffe, 2003, p. 211; Hultman, 2003/2008, p. 121).

The fact that three qualifiers (indefinite, generalizing, quantitative) are applied to one and the same pronoun need not be considered curious ('indefinite' has been interpreted as subsuming 'general', cf. Kordić, 2002, p. 49) but *man* has other, less expected, usages. The following two are characteristic of the colloquial language and are not as regularly described in reference books but they have been noted in some: *man* can be used by the speaker to refer to him or herself (all the monolingual dictionaries of Swedish listed at the end record such usage and Holmes and Hinchliffe (2003) describe it "as a polite or mildly ironical substitute for *jag* ["I"]" (p. 149)). They also describe the use of *man* instead of *du* ("you, sg."), as a "familiar term of address with a touch of ironical politeness and formality" (ibid.), as does *SAOB* (entry: *man*). Even without having mentioned the latter two uses of *man*, Ljung and Ohlander (1971/1982) warn that one has to be careful in translation and point out that the pronoun has a number of equivalents in English: one, you, we, they, people, etc. (p. 130).

---

<sup>1</sup>Translation of this, as well as of all the following quotes from Swedish or Croatian, is the author's.

### 3.2 Croatian Equivalents

A very obvious candidate for the Croatian equivalent of *man* is the substantive *čovjek* (“man”). The difference in the word class can be neglected in view of the Swedish pronoun’s etymology (developed from the substantive *man*, cf. *SAOB*: *man*) and other obvious similarities [cf. Giacalone Ramat and Sansò (2007) who focus on “indefinite man-constructions” in which “the subject position is filled by (an element deriving etymologically from) a noun meaning ‘man’”. (p. 95)]. In her study of the Croatian substantive, Kordić (2002) notes that “[s]ome nouns have a very general meaning, similar to that of pronouns” and that this is “primarily true of the noun *čovjek*” (p. 49).

Articles for the entry *man* in Swedish-Croatian dictionaries<sup>2</sup> also quote *čovjek* as an equivalent along with several others, complemented with an explanation (“indicates an unspecified person”) and examples of usage, whose Croatian translations include other equivalents than the ones quoted. In the two print editions consulted, *ljudi* (pl. form of *čovjek*, “people”) is the only other equivalent, while *Lexin online* lists *netko* (“someone”) and *neki* (“some”), with *čovjek* quoted third and *ljudi* not appearing at all. Translations of the usage examples contain, in addition to *čovjek*, the so called *se*-passive construction as an equivalent to *man*: “*man räknar med en rekordskörd i höst – računa se s rekordnom žetvom na jesen*” (*se*-passive construction is composed of the reflexive pronoun *se* and the active form of the finite verb. The Croatian sentence thus reads [A] *record harvest in [the] autumn is counted with*, rather than the Swedish source’s *One counts with a record harvest in [the] autumn*).

A cursory look at several web sites offering online Swedish-Croatian dictionaries has not identified any additional equivalents to *man*. In other words, altogether five lexicographic equivalents have been established: *čovjek*, *ljudi*, *netko*, *neki*, *se*-passive. Considering the various meanings of *man* described above, the list can hardly be exhaustive, and a good way to complete it is certainly by analyzing translation equivalents of the Swedish pronoun in a translation corpus.

## 4 Aims and Material

In view of all that has been said, a double aim was envisaged for the study: firstly, to compare the translation equivalents of *man* in a translation corpus to the list of lexicographic equivalents, assess the latter in that light and, if it proves justified, suggest changes to the list; secondly, to see whether the set of Croatian translation equivalents established in the corpus offers any other, if only preliminary, insights, in particular regarding the usage of the pronoun *man* and the nature of the equivalence established.

<sup>2</sup>Among the very limited number of Swedish-Croatian dictionaries, editions of the *Lexin* series are considered standard reference works. Two printed and one on-line edition have been consulted, in which two different versions of the article for the entry *man* have been established.

When discussing limitations of the corpus-based approach, Granger (2003) observed that “The dearth of parallel [i.e., translation] corpora accounts for the relatively large number of cross-linguistic studies which are corpus-based in the sense that they rely on authentic texts rather than introspective methods, but not in the more usual sense of computer corpus-based” (p. 23). The corpus compiled for this study is electronic but it has had to be of the “do-it-yourself kind” (cf. Mikhailov & Cooper, 2016, p. 39). It consists of six Swedish novels and their published Croatian translations (five translators) as well as 105 standard pages of Swedish non-literary texts of different types and their translations into Croatian taken from the Masters theses (the Swedish-Croatian translation part) by five graduate students of Swedish/translation stream at the University of Zagreb. The Swedish subcorpus exceeds 410,000 words and it includes 998 occurrences of the pronoun *man* in its subject form (the object form *en* and the possessive form *ens* have not been included in the analysis at this stage). The two subcorpora have been aligned, the Swedish one searched for pronominal *man* and the equivalents in the Croatian subcorpus established. Since some Swedish sentences containing *man* were left out in translation, 988 pairs of sentences have eventually been analysed. The aims set for the study have required that the analysis be both quantitative and qualitative.

## 5 Procedure and Results

The first stage of analysis has shown clearly that *man* presents a translation problem in Swedish-Croatian translation which cannot be solved by choosing from among a few obvious equivalents. On the contrary, three groups of solutions emerged based on the size of the segment for which an equivalent could be established. In the Croatian sentences that have an equivalent to *man* at the word level, 23 different equivalents appear, either pronouns (19) or nouns (4). Among the nouns, three are indeed specific nouns, namely *čovjek* (“man”), *ljudi* (“people”) and *osoba* (“person”), while the fourth stands for the category ‘noun’, i.e., different nouns that were used when the reference of the Swedish *man* got explicated in the translation, based on the context. An example is

Sw. *Man kom igång med att hurra lite och ställde sig upp (...)* (“One started (...)”)

Cro. *Gosti su počeli nazdravljati ustavši od stola (...)* (“[The] guests started to cheer having risen from the table (...)”).

The second group of translation solutions includes three verbal constructions which actually correspond to the combination of *man* and the finite verb in the source sentence. These are the already mentioned *se*-passive, as well as impersonal verbal constructions and infinitive:

Sw. *Sådant bara vet man, det går inte att förklara.* (“One just knows such (...)”)

Cro. *Takve se svari jednostavno znaju, to se ne da objasniti.* (*se*-passive; “Such things are simply known (...)”);

Sw. *Man fick vara glad åt det lilla.* (“One had to be happy (...)”)

Cro. *Treba biti zadovoljan i s malim.* (impersonal verb, subject not expressed; “[One] has to be happy also with [the] small”);

Sw. *De visste inte hur man överlevde i fiendelägret.* (“(...) how one survived (...)”)

Cro. *Ne znaju kako preživjeti u neprijateljskom kampu.* (infinitive; “They don’t know how to survive (...)”).

The third set of solutions involves equivalence only at the sentence (clause) level, with translations that either have a different subject or for some other reason need no word-equivalent to *man*. Such sentences are often constructed through modulation or use of Croatian idiomatic language:

Sw. *Jag försöker skrika, men det är svårt om man har tungan fastklistrad.* (“(...) if one has the tongue glued up”)

Cro. *Pokušavam vikati, ali to je teško kad je jezik zalijepljen.* (modulation; “I try to scream but that’s difficult when [the] tongue is glued up.”)

Sw. *Vad säger man?* (“What does one say?”)

Cro. *Je l’ tako?* (idiomatic language; “Is [that] so?”).

As the next step, the equivalents of the first kind (23 altogether) were grouped in order to match the usual description of the functions of *man*, disregarding differences that seemed irrelevant in that context. The various indefinite pronouns were thus lumped together (*svi* “all”, *svatko* “everybody”, *nitko* “nobody”, *mnogi* “the many”, *tko* “who”) to match the “indefinite” or “indefinite quantitative” description of *man*, with the exception of the pronouns that are quoted in Croatian dictionaries, which were kept as a separate group (*neki* “some”, *netko* “somebody”). At first glance it might seem that the same could be done with personal pronouns of the same person since the number does not change the relevant functions (e.g., speaker/addressee). That is, however, not the case both because of the function of *man* (replacing *jag* and *du*, rather than the plural pronouns of the same person) and because of the fact that some Croatian personal pronouns have, along with their prototypical referential function, an impersonal use and it is exactly the number that is relevant in that respect (only 2nd person singular and 1st and 3rd persons plural can have that function, cf. Silić & Pranjković, 2005, p. 318). Personal pronouns posed another dilemma in that they sometimes appear as an explicit one-word equivalent for *man* but are more commonly implicit in sentences with no explicit subject and the grammatical person marked in the finite verb (the Croatian syntax allows for unexpressed subjects). The corpus thus contains e.g.,

Sw. *Får man prova?* (“May one try?”)

Cro. “Smijem i ja probati?” (“May I try, too?”, subject “I” expressed)

as well as

Sw. *Man måste tänka på pengarna!* (“One must think about the money!”)

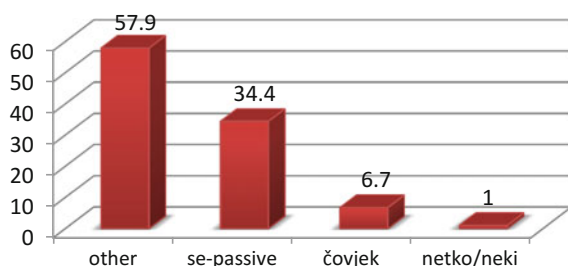
Cro. *Moram misliti na novac.* (“[I] must<sub>1st sg</sub> think about (...)”, unexpressed subject, the verb form indicating 1st person singular).

Since it is the meaning of the person and number that is essential here rather than the way in which it is expressed, examples of both kinds were taken to belong to the same category. Eventually, eighteen different categories of translation equivalents were thus obtained and their respective shares in the overall number (i.e., 988) calculated.

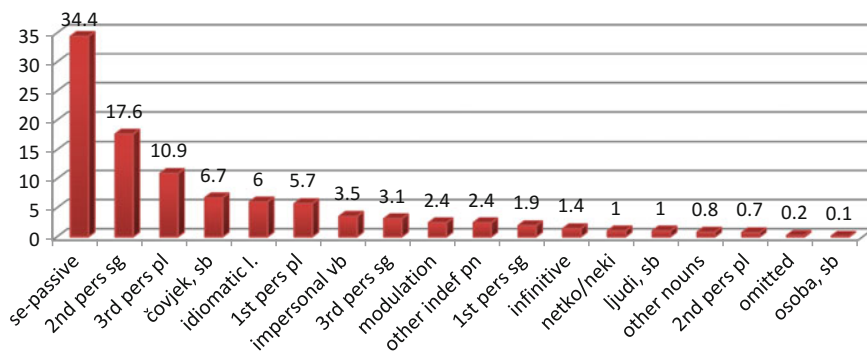
The first point of interest was to see how common the equivalents established in the dictionaries were within the set of translation equivalents. As can be seen in the next graph (Fig. 1), taken together, they account for less than a half of all the translation solutions (42.1%). In the 988 translated sentences, the *se*-passive construction has been used 340 times (34.4%), the word-equivalent *čovjek* 66 times (6.7%), pronouns *netko* or *neki* altogether 10 times (1%). In 572 sentences (57.9%) equivalence has been established in a way not mentioned in the dictionaries.

In the next step the heterogeneous category 'other' was broken up and the share in the corpus of each of the eighteen categories mentioned above was established (Fig. 2).

At this stage it has become clear that the Croatian *se*-passive construction is the most common translation solution in the corpus (340 occurrences). Following it are the 2nd person singular markers (personal pronoun *ti* or finite verb forms,



**Fig. 1** Share (%) of lexicographic equivalents of *man* (*se*-passive, *čovjek*, *netko/neki*) versus other translation solutions in the corpus



**Fig. 2** Share (%) of the various translation equivalents in the corpus



174 occurrences) and the 3rd person plural markers (pronouns *oni, one, ona* or verb forms, 108 occurrences). The next three categories all have a share of between 5 and 7% in this corpus: *čovjek* (66), equivalence established through the use of idiomatic language (59), and the markers of 1st person plural (personal pronoun *mi* or appropriate verb forms, 56 altogether). Following these is the category ‘impersonal verb’ (35), in which an impersonal verb is used in translation, with the subject neither expressed nor implied. Slightly less frequent in this corpus is the use of the 3rd person singular markers (pronouns *on, ona, ona*/verb forms, 31). Two very different types of solution occur 24 times each: equivalence achieved at the sentence/clause level through modulation and by using an indefinite pronoun other than *netko* or *neki* (if the two groups of indefinite pronouns were not kept separate, which has been done here only in order to enable an assessment of lexicographic equivalents, the overall share of indefinite pronouns in the corpus would be 3.4%). The markers of the 1st person singular (personal pronoun *ja* or the appropriate verb forms) occur 19 times as translation equivalents and are followed by the infinitive verb form (14) and the two remaining categories involving lexicographic equivalents i.e., the indefinite pronouns *netko or neki* and the noun *ljudi* with 10 occurrences each. The remaining four types of solutions have been used less than 10 times.

## 6 Discussion

A comparison of the results of the quantitative analysis and the articles for the entry *man* in the consulted dictionaries shows that the latter do not fully reflect the variety of uses of the Swedish pronoun. Notwithstanding the fact that many different circumstances determine the lexicographic presentation, some changes in what has been established as the standard set of lexicographic equivalents and explanations are warranted. The following seems relevant:

- the *se*-passive construction has proved to be the most frequently used translation solution, almost twice as common as the one following it, while its dictionary status is only minor (not specifically referred to at all, just appearing in translations of the usage examples);
- while the category ‘2nd person singular’ has the second largest share in the corpus, it is not mentioned in the dictionaries at all. As was indicated above, its common use in Croatian translations has to do with the fact that it is not only used in reference to the addressee but it can also express impersonal meaning (cf. Silić & Pranjeković, 2005, p. 318), as in

Sw. Rädsla är något **man** lever med ensam. (“Fear is something one lives with.”)

Cro. Strah je nešto s čime živiš sam. (“Fear is something with which [you] live<sub>2nd sg</sub> alone.”)

The impersonal use of the 2nd person singular, as well as of the 1st and the 3rd person plural, may be considered ‘covered’ by the dictionary explanation “indicates

an unspecified person”. However, the corpus also includes examples of the 2nd person with specific reference to the addressee, as in

Sw. “Hörrudu din lille fjant!” fräste han. “Plötsligt är **man** så jävla kaxig! Akta dig ditt lilla äckel!” (“All of a sudden one is (...)”)

Cro. - Čuj, ti mali glupane! - zaurlao je. Odjednom **si** tako bezobrazan! Samo se pazi, govнару mali! (“Listen, you little jerk!, he shouted. All of a sudden [you] are<sub>2nd sg</sub> so cheeky.”; both texts are a direct address).

The same is true of the 3rd and the 1st person plural and none of that is accounted for in the dictionaries, not even with usage examples.

- the possibility for *man* to refer to the speaker is not mentioned in the dictionaries either, while it is acknowledged in Swedish reference books and the category ‘1st person singular’ appears as the chosen translation solution 19 times in the corpus. In other words, no element in the dictionary articles provides an explanation for sentences such as the following:

Sw. - Får **man** vara med? (“May one participate/join in?”)

Cro. - Mogu li igrati s vama? (“May<sub>1st sg</sub> [I] play with you?”)

- the noun *čovjek* is rightly quoted as a lexicographic equivalent but this is not always done in a way which would make it obvious that it is its generalizing use that corresponds to *man*;
- the pronoun *netko* is one of the two most common indefinite pronouns among the translation equivalents of *man*. The other one is, however, not the pronoun quoted in the dictionaries (*neki*, only one occurrence in the corpus) but *nitko* (“nobody”) with seven occurrences.

With regard to the second aim of this study—to check whether the compiled translation corpus can offer any insights into other phenomena, linguistic, contrastive linguistic or those that might be of interest for translation studies—several observations can be made. As has already been stated, multiple interpretations of a linguistic unit, in this case *man*, contained in a translation corpus expose the unit’s various meanings and uses. In the light of the results obtained in this study, it would seem interesting to check how often the ‘indefinite pronoun’ *man* actually has a known referent and whether that circumstance should be made more prominent in its description. Such a usage prompts translations categorized here as ‘other nouns’ or those grammatical persons that normally do not imply an impersonal meaning, but it is also manifest in many cases where modulation has been applied as well as in a number of translations involving the three grammatical persons that may but need not be used impersonally. It would seem that the overall share of such translation solutions is not negligible. On the other hand, one might easily come to an erroneous conclusion if the translation equivalents were taken to necessarily be truly identical to the source unit in focus. The impression obtained through this study suggests that Croatian equivalents of *man* often do not preserve its distancing function and that they resolve the (deliberate) ambiguity with regard to its referent. Any conclusions about the use of *man* would therefore have to be substantiated in a monolingual corpus or in another way.

The comments on the interpretation of the referent and the equivalence established are illustrated by the following:

Sw. Och bjuder **man** Pia måste **man** bjuda Eva-Lena, nästan skriker Juha.

(“And if one invites Pia, one must invite Eva-Lena, Juha almost shouts.”)

Cro. A ako pozovemo Piju, onda moramo pozvati i Evu-Lenu – Juha sad gotovo viče (...)

(“And if [we] invite<sub>1st pl</sub> Pia, then [we] must<sub>1st pl</sub> also invite Eva-Lena – Juha is almost shouting now”, “we” refers to Juha and his friend with whom he is planning a party, no generalized interpretation is possible)

Sw. - Så ska **man** få det kastat i ansiktet också, sa modern. (“So one shall get it/that thrown in the face too, said the mother.”)

Cro. “Sad **me** i za to kriviš“, rekla je majka. („Now you blame me for that too“).

Further research could also try to establish whether the disambiguating translation equivalents in the corpus may be taken as resulting from the translator’s choice only and thus indicating that translators find it more important to produce a referentially unambiguous translation than to preserve the generalizing, distancing or a similar quality.

## 7 Conclusion

In line with the view of translation corpora as a valuable research tool in both Contrastive Linguistics and Translation Studies, the study presented here identifies some potential uses of a translation corpus of the do-it-yourself kind, such as is still necessary for many language pairs. The Swedish-Croatian translation corpus has been analysed in order to get a more comprehensive understanding of the various ways in which the Swedish pronoun *man* can be rendered in Croatian and to assess the standard lexicographic presentations of the pronoun in bilingual dictionaries in that light. The qualitative analysis required towards that goal has indicated that the corpus could be fruitfully used to pursue other research topics, as different as e.g., relating individual Croatian equivalents to a particular function of *man* on the one hand and assessing translators’ preferences or the nature of translation equivalence on the other.

## References

- Aijmer, K. (2007). The meaning and functions of the Swedish discourse marker *alltså*—Evidence from translation corpora. *Catalan Journal of Linguistics*, 6, 31–59.
- Altenberg, B., & Aijmer, K. (2000). The English-Swedish parallel corpus: A resource for contrastive research and translation studies. In Ch. Mair & M. Hundt (Eds.), *Corpus linguistics and linguistic theory. Papers from the 20th International Conference on English Language Research on Computerized Corpora. (ICAME 20) Freiburg im Breisgau 1999*. Amsterdam, Atlanta, GA: Rodopi.

- Altenberg, B., & Granger, S. (2002). Recent trends in cross-linguistic lexical studies. In B. Altenberg & S. Granger (Eds.), *Lexis in Contrast: Corpus-based Approaches*. Amsterdam, Philadelphia: John Benjamins. 3–48.
- Baker, M. (1993). Corpus linguistics and translation studies: Implications and applications. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology. In honour of John Sinclair*. Amsterdam, Philadelphia: John Benjamins.
- Baker, M. (1995). Corpora in translation studies: An overview and some suggestions for future research. *Target*, 7(2), 223–243.
- Bolander, M. (2005). *Funktionell svensk grammatik*. Stockholm: Liber AB.
- Filipović, R. (1973). The use of a corpus in contrastive studies. *Studia Romanica et Anglicae Zagrabiensia*, 33–36, 489–500.
- Gellerstam, M. (1986). Translationese in Swedish novels translated from English. In L. Wollin & H. Lindquist (Eds.), *Lund Studies in English Vol. 75: Translation studies in Scandinavia* (pp. 88–95). Lund: CWK Gleerup.
- Giacalone Ramat, A., & Sansò, A. (2007). The spread and decline of indefinite man-constructions in European languages: An areal perspective. In P. Ramat & E. Roma (Eds.), *Europe and the mediterranean as linguistic areas. Convergencies from a historical and typological perspective* (pp. 95–131). Amsterdam, Philadelphia: John Benjamins. <https://doi.org/10.1075/slcs.88.07.gia>.
- Gómez-González, M. Á., & Doval-Suárez, S. M. (2005). On contrastive linguistics; Trends, challenges and problems. In C. S. Butler, M. Á. Gómez González, & S. Doval Suárez. (Eds.), *The dynamics of language use: functional and contrastive perspectives* (pp. 19–44). Amsterdam, Philadelphia: John Benjamins.
- Granger, S. (2003). The corpus approach: a common way forward for contrastive linguistics and translation studies? In S. Granger, J. Lerot, & S. Petch-Tyson (Eds.), *Corpus-based approaches to contrastive linguistics and translation studies* (pp. 17–29). Amsterdam, New York, NY: Rodopi.
- Hoey, M. & Houghton, D. (2001) Contrastive analysis and translation. In M. Baker (Ed.), *Routledge encyclopedia of translation studies* (pp. 45–50). London, New York, NY: Routledge.
- Holmes, Ph, & Hinchliffe, I. (2003). *Swedish. A comprehensive grammar* (2nd. ed.). London, New York: Routledge.
- Hultman, T. G. (2003/2008). *Svenska Akademiens språklära*. Stockholm: Svenska Akademien.
- Ivir, V. (1997). Formal/contrastive correspondence and translation equivalence. *Studia Romanica et Anglicae Zagrabiensia*, 42, 167–180.
- Josefsson, G. (2009/2011). *Svensk universitetsgrammatik för nybörjare* (2nd ed.). Lund: Studentlitteratur.
- Koridć, S. (2002). *Riječi na granici punoznačnosti*. Zagreb: Hrvatska sveučilišna naklada.
- Lan, L. (2014). Corpus. In S.-w. Chan (Ed.), *The Routledge encyclopedia of translation technology* (pp. 465–479). London, New York, NY: Routledge.
- Lewandowska-Tomaszczyk, B. (2012a). Explicit and tacit. An interplay of the quantitative and qualitative approaches to translation. In M. Oakes & M. Ji (Eds.), *Quantitative Methods in Corpus-Based Translation Studies: A practical guide* (pp. 3–34). Amsterdam – Philadelphia: John Benjamins.
- Lewandowska-Tomaszczyk, B. (2012b). Cognitive corpus studies: A new qualitative & quantitative agenda for contrasting languages. *Connexion - A Journal of Humanities and Social Sciences*. Mae Fah Luang University, Thaliand. 29–63. <https://doi.org/10.14456/connexion.2012.2>.
- Ljung, M. & Ohlander, S. (1971/1982). *Allmän grammatik*. Malmö: Liber Förlag.
- META: journal des traducteurs/Translators' Journal*. 43(4). (1998). Montréal: Les Presses de l'Université de Montréal.
- Mikhailov, M., & Cooper, R. (2016). *Corpus linguistics for translation and contrastive studies*. London, New York: Routledge.
- Ramón García, N. (2002). Contrastive linguistics and translation studies interconnected: The corpus-based approach. *Linguistica Antwerpiensia*, 1, 393–406.
- Silić, J. & Pranjković, I. (2005). *Gramatika hrvatskoga jezika*. Zagreb: Školska knjiga.
- Toury, G. (1995). *Descriptive translation studies and beyond*. Amsterdam, Philadelphia: John Benjamins.

## Dictionaries

- Lexin online, svenska – kroatiska*, a cooperation between Institutet för språk folkminnen and Kungliga tekniska högskolan. Retrieved from [http://lexin.nada.kth.se/lexin/#about=2;searchinfo=both\\_swe\\_hrv\\_man](http://lexin.nada.kth.se/lexin/#about=2;searchinfo=both_swe_hrv_man). *Lexin. Švedsko-hrvatski rječnik/ Svensk-kroatiskt lexikon*. (1994). Zagreb: Naklada Dominović, Stockholm: Norstedts Förlag.
- Lexin. Švedsko-hrvatski rječnik/ Svensk-kroatiskt lexikon*. (2001). Stockholm: Statens skolverk.
- Nationalencyklopedins ordbok*, prepared at Språkdata, Göteborgs universitet. (1995/2000). Höganäs: Bokförlaget Bra Böcker.
- SAOB online, Svenska Akademiens ordbok*. Retrieved from [https://www.saob.se/artikel/?unik=M\\_0157-0022.cGb&pz=5](https://www.saob.se/artikel/?unik=M_0157-0022.cGb&pz=5).
- Svensk ordbok*, prepared at Språkdata, Göteborgs universitet. (1986). Språkdata and Esselte studium.
- Svensk ordbok utgiven av Svenska Akademien, M-Ö*. prepared at Redaktionen för Svenska Akademiens samtidsordböcker and Lexikaliska institutet, Institutionen för svenska språket, Göteborgs universitet. (2009). Stockholm: Svenska Akademien, Norstedts.

## Corpus Sources

### I. Published:

- Fredriksson, M. (1997). *Enligt Maria Magdalena*. Stockholm: Wahlström&Widstrand.
- Fredriksson, M. (1999). *Prema Mariji Magdaleni* (S. Žuljević, Trans.). Zagreb: VBZ.
- Gardell, J. (1992). *En komikers uppväxt*. Stockholm: Norstedts.
- Gardell, J. (2004). *Odrastanje komičara* (Ž. Černok, Trans.). Zagreb: AGM.
- Lindgren, A. *Alla vi barn i Bullerbyn*. Stockholm: Rabén & Sjögren.
- Lindgren, A. (2003). *Mi, djeca Graje Male* (G. Antunović, Trans.). Golden marketing.
- Marklund, L. (2000). *Paradiset*. Piratförlaget.
- Marklund, L. (2005). *Raj* (Ž. Černok, Trans.). Zagreb: Algoritam.
- Martinson, H. (1994). *Nässlorna blomma. Alla Tiders Klassiker*. Stockholm: Rabén & Sjögren.
- Martinson, H. (1987). *Kad su cvjetale koprive* (M. Rumac, Trans.). Zagreb: Liber.
- Niemi, M. (2000). *Populärmusik från Vittula*. Pan. Stockholm: Norstedts förlag.
- Niemi, M. (2005). *Pop glazba iz Vittule* (D. Maček, Trans.). Zagreb: AGM.
- II. Swedish-Croatian translation part of the following diploma theses at the Scandinavian Section, Faculty of Humanities and Social Sciences, University of Zagreb:
- Grubješić, I. (2015). *Översättning av passiv från svenska till kroatiska*.
- Luketić, P. (2013). *Översättning av institutionsnamn mellan kroatiska och svenska*.
- Profeta, S. (2012). *Meningslängd i olika texttyper vid översättning från svenska till kroatiska*.
- Veršić, V. (2014). *Ortnamn i svensk-kroatisk översättning*.
- Žic, L. (2014). *Översättning av neologismer och verkspecifika termer*.
- Altenberg, B., & Granger, S. (2002). Recent trends in cross-linguistic lexical studies. In B. Altenberg & S. Granger (Eds.), *Lexis in Contrast: Corpus-based Approaches*. Amsterdam, Philadelphia: John Benjamins. 3-48.

**Goranka Antunović** is Associate Professor at the Faculty of Humanities and Social Sciences, University of Zagreb, where she is the head of the Scandinavian Studies Section. Her main research interests lie in the field of translation studies (translation norms, process studies, currently focusing on post-editing and translator training) as well as in the field of contrastive linguistics (Swedish-Croatian corpus-based contrastive studies in particular). Her titles include “The

effect of interpreting experience on distance dynamics: Testing the literal translation hypothesis” (a chapter written together with Nataša Pavlović in Ehrensberger-Dow, M.; Englund Dimitrova, B.; Hubscher-Davidson, S.; Norberg, U. *Describing Cognitive Processes in Translation: Acts and events*. Amsterdam: John Benjamins. 2015). She is the author of *A Dictionary of Anglicisms in Swedish* (Zagreb, 1999) and the editor of several volumes, including *Glimpses of the North: Discovering Scandinavia and Scandinavian Studies*. (Zagreb: Srednja Europa. 2014) and *Language Varieties Between Norms and Attitudes—South Slavic Perspectives. Proceedings from the 2013 CALS Conference* (together with A. Peti-Stantić and M. Stanojević. Frankfurt am Main: Peter Lang). G. Antunović is a member of the editorial board of *Hieronymus, Journal of Translation Studies and Terminology* (<http://www.ffzg.unizg.hr/hieronymus/>).