THEY ARE KIND OF/SORT OF SIMILAR: A PARALLEL CORPUS-BASED ANALYSIS OF ENGLISH *KIND OF* AND *SORT OF* AND THEIR LITHUANIAN CORRESPONDENCES

Audronė Šolienė

Vilniaus universitetas
Universiteto g. 5
LT-01131 Vilnius, Lietuva
El. p. audrone.soliene@flf.vu.lt

1. Introduction

In Present-day English, there are three main nouns expressing the general meaning of 'type': *type of*, *kind of* and *sort of*. Traditionally, the type noun constructions are defined as nominal expressions dealing with (sub)categorization (Brems 2011, 2), or in other words, as nominal phrases used to ascribe a lexical item to a particular group or type sharing exact or similar features, i.e. dealing with taxonomic relationships, as in the example "a (special) sort of rose" (Traugott 2008, 226). The definition, however, applies only to the lexical meaning of these nouns. It must be admitted that, throughout the years, *kind of* and *sort of*, via the grammaticalization process, have gone through the pragmatic-semantic process of subjectification, which, according to Traugott (1995, 32), is:

> a gradient phenomenon, whereby forms and constructions that at first express primarily concrete, lexical, and objective meanings come through repeated use in local syntactic contexts to serve increasingly abstract, pragmatic, interpersonal, and speaker-based functions.

Thus, the term *type noun* in this paper encompasses not only the primary lexical meaning, but is also used to account for other pragmatically charged uses of the items in question.

The English *kind of* and *sort of* have received a lot of attention in monolingual studies.[1] It has been proved that *kind of* and *sort of* are extremely multifunctional (i.e. they may indicate a type or species, convey a vague reference, signal inadequate word choice, express imprecision or hesitation, diminish intensity, protect face, or fill in pauses, etc.). They have been investigated under different headers (e.g. species/type nouns, pragmatic markers, pragmatic particles, discourse markers, vagueness or fuzziness markers, hedges, downtoners, stance adverbials, adverbs of degree, etc.) (see Kay 1984; Holmes 1988; Aijmer 2002; Gries, David 2007; Davidse et al. 2008; Brems, Davidse 2010; Fetzer 2010; Margerie 2010; Kirk 2015, inter alia).

Though there have been a number of contrastive corpus-based studies of different discourse markers (DMs) carried out (see Aijmer, Simon-Vandenbergen 2003; Lewis 2006; Johansson 2007; Degand 2009; Beeching, Detges 2014; Furkó 2014 among others), which have resulted in prolific cross-linguistic and translation research output on English, French, Spanish, German, Dutch, Swedish, Norwegian, Hungarian and other languages, contrastive parallel corpus-based studies of *kind of* and *sort of* in particular, involving two or more language sets, are extremely scarce. To my knowledge, one of the exceptions is Janebovà and Martinkovà's paper (2017), which accounts for their research of *kind of* and *sort of* and their Czech correspondences in the English-Czech section of InterCorp. Their results demonstrated that the type nouns not only carry several distinct functions but also have an array of different translations in Czech. However, studies comparing *kind of* and *sort of* and their Lithuanian equivalents are non-existent.

This cross-linguistic research sets out to describe the quantitative and qualitative distribution of the English *kind of* and sort *of*, to determine the translational correspondences (TCs) of the two linguistic units in Lithuanian, as well as to reveal their functional diversity in

---

[1] This paper leaves the constructions with *type* out of the discussion since, as claimed by Brems and Davidse (2010, 182), *type* is a later addition to the set of the types nouns and it "<...> developed at a different rate than *sort* and *kind*; qualifying uses with *type*, for instance, are only just starting to be attested."

terms of (inter)subjectivity (Traugott 2010), and to investigate how Lithuanian correspondences correlate with the function that *kind of* and *sort of* perform in original and translated fiction texts:[2]

(1)   EN-orig: *Usually Hatsumomo liked to drink* **a special kind of sake** *called amakuchi – which was very light and sweet.*

LT-trans: *Paprastai Hacumomo mėgdavo gerti* **ypatingos rūšies sakę**, *vadinimą amakuči. Ji buvo labai silpna ir saldi.*

(2)   EN-orig: *"You know* **what kind of chap** *I was."*

LT-trans: – *Tu ir pats žinai,* **kas aš per žmogus.**

(3)   EN-orig: *"I sense you've heard of antimatter, Mr. Langdon?"<...> Langdon looked up. He felt suddenly dumb. "Yes. Well ...* **Sort of**.*"*

LT-trans: – *Nujaučiu, jūs esate girdėjęs apie antimedžiagą, pone Lengdonai? <...> Lengdonas pažvelgė į ją. Kažkodėl jam užkando žadą. – Taip... Žinoma...* **Kažkas tokio.**

The three examples above illustrate the range of meaning (from objective to (inter)subjective (Traugott 2010)) *kind of* and *sort of* may display: moving from a marker of a clear taxonomic relationship (1) to a more descriptive post-determiner use (2), and finally to a free adjunct functioning in the interpersonal domain (3).

---

[2] Examples whose source origin is ParaCorp$_{EN \to LT \to EN}$ (see Section "Data and Methods") are labelled with tags indicating the source and target languages: EN-orig, EN-trans, LT-orig and LT-trans. Other examples provided are marked according to general referencing rules.

2.   From type nouns to discourse markers

Scholars seem to disagree on the exact number of the constructions the type nouns can appear in. However, traditionally there is a unanimous agreement that *kind of* and *sort of* may be used noun–phrase (NP) internally (i.e. nominal use) or noun–phrase externally, i.e. they may be exploited to modify adjectives, verbs and adverbs, prepositional phrases or function as completely independent units (i.e. free adjuncts) modifying whole propositions.

When *kind of* and *sort of* are used NP-internally, the first construction they appear in is the **binominal** one (Denison 2005; Traugott 2008), also known as lexical–head use (Davidse et al. 2008, 147), in which the type noun has a generic reference, is used in hyponymy statements, and denotes a specific subclass or, as in example (1) above and examples (4) and (5):

(4)   EN–orig: *They fill you with bubbles, and the bubbles are full of **a special kind of gas**, and this gas is so terrifically lifting that it lifts you right off the ground just like a balloon* <...>.

(5)   EN–orig: *He untied it and took out a short, fat scroll, which he began to unroll. It was cracked with age and showed-in miniature-brilliantly colored scenes of the Imperial court. If you've ever seen **this sort of scroll**, you'll know that you can unroll it all the way across a room* <...>.

Here the type noun is the head and the *of–phrase* is its modifier (Brems, Davidse 2010).[3] The binominal *kind of* and *sort of* may "either introduce a (potential) discourse topic or refer back to an existing discourse topic" (Keizer 2007, 155-156). The type noun can be both singular and

---

[3] Following Keizer (2007), Kolyaseva and Davidse (2016) maintain that the *of–phrase* here functions as a complement of the head noun; however, this theoretical issue of complementation and modification falls out of the scope of the present paper.

plural in number and all determiners are freely available in the pre-modification field of the phrase (Traugott 2008, 234; Denison 2005). They are not restricted to affirmative environments but can also appear in questions.

The second construction of the nominal *kind of* and *sort of* is the **qualifying** construction (Denison 2005), or degree modifier (Traugott 2008, 226-227), which "is the only one that is unanimously recognized besides the head use" (Davidse et al. 2008, 156). The qualifying construction, in contrast to the binominal construction, does not carry a generic reference; here the type noun functions as a discourse marker and hedge and "conveys that the description is only approximate" (Davidse et al. 2008, 157) or, according to Traugott (2008, 227), "[p]ragmatically such phrases cast doubt on the accuracy of the description", e.g.:

(6)   EN-orig: *"I think that was **a sort of** joke," Mameha said.*

In this construction the type nouns are modifiers having the meaning of *more or less* or *somewhat* and "NP2 is perspectivized as in the foreground" (Traugott (2008, 226-27). Here the type noun normally occurs in its singular form and is preceded by an indefinite determiner or no determiner at all.

Alongside the two above-mentioned NP-internal constructions, *kind of* and *sort of* can be used NP-externally, i.e. they can appear in **adverbial** constructions where they function as unquestionable discourse markers (Kay 1984, 158; Dehé, Stathi 2016, 917). They can modify adjectives (7), verbs (8), adverbs (9), prepositional phrases[4] and entire clauses (10) or function as stand-alone adjuncts (11), for instance:

(7)   EN-orig: *Most of it was **kind of silly**.*

---

[4] Cases where the type nouns modified PPs were not attested in the present analysis, possibly due to the limitation of the data sources.

(8)   EN-orig: *I **sort of like** you.*

(9)   EN-orig: *She **sort of accidentally** made it sound like I was trying to kill myself or something.*

(10)   EN-orig: *That's **kind of how they got into this mess in the first place**.*

(11)   EN-orig: *At least it smelled clean. **Sort of.***

The non-NP internal construction fulfils numerous "discourse functions such as self-repair, hedging strong opinions, establishing common ground between interlocutors, politeness, and so forth" (Dehé, Stathi 2016, 918).

In addition to the binominal, qualifying and adverbial constructions, a number of **other constructions** are recognised in the literature. One of them is the post-determiner/complex determiner construction where *kind of* and *sort of* also feature noun-phrase internally. The question whether the type noun is the head of the NP or not remains to be open. Denison (2005) seems not to offer a straightforward answer whether this is a distinct construction at all. Keizer (2007) considers this construction as a sub-group of the binominal construction; however, Davidse et al. (2008), Brems and Davidse (2010), and Janebovà and Martinkovà (2017) treat this construction as a separate one as "[t]he post-determiner construction is the result of reanalysis of the binominal construction" (Brems, Davidse 2010, 181); it no longer has a generic reference, but has a mainly textual function and displays phoric (anaphoric and cataphoric) relations, so it should be treated as a distinct category, i.e. instead of being strictly taxonomic, these units have their meaning expanded to put forward the idea of a description or indication that something is of a certain nature:

(12)   EN-orig: *Unfortunately it was **the kind of job** that only keeps hands busy.*

Another pattern proposed by Davidse et al. (2008) is a quantifying construction. This construction is quite special, because, due to lexicalization, the type nouns in plural appear in fixed phrases with *all* and carry the meaning of 'many/much':

(13)   EN-orig: *Fache had posited **all kinds of explanations** tonight to explain Sophie's odd behavior <...>.*

In addition to quantifying uses with *all*, negative totality quantifiers as *no* or (*not*) *any* also fall under this category.

Finally, the NP-internal *kind of* and *sort of* can function as modifiers. The main characteristic of the modifier use of the type nouns is the use of an adjective or a noun that is more related to N2 than the type noun itself, for instance:

(14)   *It's a **cool quirky kind of** song.* (Davidse et al. 2008, 147)

Here *kind of* and *sort of* have a metalinguistic value, indicating that the lexical items preceding them "have to be interpreted as ad hoc, often very creative, classifiers" (Brems, Davidse 2010, 188). They also act pragmatically, signalling that "the description is only approximative" (ibid.).

In this paper a distinction is made between the propositional use of *kind of* and *sort of* expressing 'a type of' and their discourse marker use. All the cases of *sort of* and *kind of* where they do not mean 'a type of' are treated as having a discourse qualifying value, i.e. functioning as discourse markers (Aijmer 2002, 178). Discourse markers are perceived as having "by definition a discourse function, which entails indexing the utterances to the surrounding discourse, both in terms of structuring the ongoing discourse and in terms of signalling to the addressee how he/she should interpret the speaker's stance" (Wichmann et al. 2010, 107).

3. Data and Methods

The present paper deals with a quantitative and qualitative contrastive corpus-based analysis drawing on the data extracted from a self-compiled bidirectional parallel corpus – ParaCorp$_{EN \to LT \to EN}$ (Šolienė 2013). The corpus is compiled according to the model of the English-Norwegian Parallel Corpus (Johansson 2007). The ParaCorp$_{EN \to LT \to EN}$ comprises original English (British and American) fiction texts and their translations into Lithuanian and original Lithuanian fiction texts and their translations into English. The choice of the data was determined by several circumstances. Firstly, fiction embodies a wide spectrum of linguistic output and style, including features of spoken discourse, with which *kind of* and *sort of* are generally associated (Biber et al. 1999, 869). Secondly, such a corpus design is advantageous as it offers different directions of linguistic comparison and can be used both as a parallel corpus and a comparable corpus (Johansson 2007, 11). The corpus consists of about 5 million running tokens (see Table 1).

Table 1. The size of two sub-corpora *ParaCorp$_{EN \to LT}$* and *ParaCorp$_{LT \to EN}$*

	Original	Translation	Total
ParaCorp$_{EN \to LT}$	1,983,266	1,541,038	3,524,304
ParaCorp$_{LT \to EN}$	608,426	788,897	1,397,323

In order to generate concordance lines with the type nouns *kind of* and *sort of* as well as their translations, the multilingual concordancer ParaConc (Barlow 1995) was used. The corpus output files in the plain text format were imported into an Excel spreadsheet for further analysis. The search was bidirectional: first, *kind of* and *sort of* were searched as nodes in the EN-orig → LT-trans direction; second, they were entered as nodes in the EN-trans → LT-orig search direction. Table 2 shows the raw frequency of the items under investigation:

Table 2. The raw frequency of *kind of* and *sort of* in ParaCorp$_{EN \to LT \to EN}$

	EN-orig ➔ LT-trans	LT-orig ➔ EN-trans
	#	#
kind of	537	232
sort of	579	122

It is important to note that the cases where *kind of* was used as an adjective in the structure *kind of somebody to + verb* were discarded from the analysis as a first step before any calculation. Based on the raw frequency count, it was decided to take 200 randomized hits from the EN-orig texts and 100 randomized hits from the EN-trans texts of each *kind of* and *sort of* from both sub-corpora for further functional analysis. All in all, there were 600 cases of the use of *kind of* and *sort of* analysed. The randomization procedure was performed using the RAND function available in the Microsoft Excel spreadsheet – each occurrence of the node in question was assigned with a random number from 0 to 1. Then the numbers were ranked from the lowest to the highest and the first 200 and 100 hits were selected for the analysis.

Though the data extraction and randomization were automatic, the qualitative analysis of the concordances was carried out manually. Firstly, the uses of the type nouns were divided into two groups depending on whether they function NP-internally or NP-externally. The NP-internal cases were grouped into the ones expressing 'type of' and the ones functioning as DMs. All of the non-internal cases were categorised according to the item they modify or act as a stand-alone adjunct. Finally, their translational correspondences as well as different functions were considered.

On the basis of the empirical data derived from a parallel corpus we can establish translation paradigms (Johansson 2007, 23) or semantic mirrors (Dyvik 2004, 311) which contain translational correspondences. First, we can divide TCs according to the direction of translation. Then having analysed the TCs of a linguistic unit in question in terms of expression, we can single out *zero* and *overt* correspondences. The latter fall into two groups:

congruent (*direct*) and *non-congruent* (*divergent*) correspondences. Congruent correspondence is the case when a linguistic item is translated into another language keeping the same part of speech (form and meaning); there is no system shift (Johansson, 2007, 24), for example:

(15)   EN-orig: *There was **this new kind of cancer** that was getting young men.*

LT-trans: *Dabar atsiradusi **nauja vėžio rūšis**, kuria serga jauni vyrai.*

Divergent correspondences, on the other hand, are more interesting for contrastive studies, as they "serve[...] as a means of uncovering differences where they may be unexpected" (Johansson 2007, 25). These are the cases when, from the perspective of meaning, a greater or lesser part of meaning of the SL text is compensated in the TL by various means, usually not listed in dictionaries as traditional prototypical equivalents (Usonienė 2006, 101), for instance:

(16)   EN-orig: *He was determined to make **this some kind of** double date, apparently.*

LT-trans: *Aišku, buvo pasiryžęs padaryti **tarsi kokį** dvigubą pasimatymą.*

Zero correspondences are cases where there is no clearly identifiable semantic counterpart in the TL or SL text (Johansson 2007, 26). Sometimes translators decide to omit a word or phrase in translation (*omission*), as in (17), or add some new shades of meaning in TL texts that were absent in the original (*addition*), as in (18):

(17)   EN-orig: *And **the kind of medicine** Peeta needs would have been at a premium from the beginning.*

LT-trans: *O **ø vaistai**, kurių reikia Pitui, jau iš pradžių kainavo labai daug.*

(18)   LT-orig: *Galėčiau tai pavadinti ø drama, kurioje asmeninei patirčiai tenka ne menkiausias vaidmuo.*

EN-trans: *I could even say it is **a sort of drama**, one in which my personal experiences have ended up playing a not insignificant role.*

The Lithuanian TCs of *kind of* and *sort of* were categorised according to the three above-mentioned types.

4.   Results

This section presents the quantitative and qualitative findings of the analysis. It starts with the analysis of the quantitative distribution of *kind of* and *sort of* (Section 4.1). The subsequent sub-sections further elaborate on both quantitative and qualitative findings.

4.1. Quantitative distribution of *kind of* and *sort of*

The normalized frequency (per 10,000 words) of *kind of* and *sort of* in both the original and translated English texts is presented in Table 3. The table gives all the relevant occurrences of *kind of* and *sort of* in the corpus.

Table 3. Raw and normalized frequencies of the type nouns in ParaCorp$_{EN \rightarrow LT \rightarrow EN}$

	EN-orig → LT-trans		EN-trans → LT-orig	
	#	f/10,000	#	f/10,000
kind of	537	**2.7**	232	**2.9**
sort of	579	**2.9**	122	**1.6**

Though the frequencies seem to be more or less levelled out across the different sub-corpora, the only discrepancy that strikes the eye is the overuse of *kind of* in the EN-trans texts in

comparison with *sort of*. The reason behind this might be the fact that the majority of the translators in ParaCorp$_{EN→LT→EN}$ are American English-based bilinguals, and some studies proved that *kind of* is much more frequent in American English than in British English (see Biber et al. 1999, 870; Quirk et al. 1985, 598).

4.1.1. The NP-external use of *kind of* and *sort of*

The non-internal use of the items under study turned out to be extremely infrequent. Out of 400 cases (200 of *kind of* and 200 of *sort of*) where the type nouns appeared in the EN-orig texts, only 18 cases of *kind of* and 35 cases of *sort of* were non-internal, which makes only 13.3 % of all of the concordances analysed (53 cases in total). Correspondingly, the NP-internal use makes up 86.7 %. Table 4 shows the parts of speech the type nouns have scope over.

Table 4. The raw frequency and percentage of non-internal use in the EN-orig texts

	Verb		**Adjective**		**Adverb**		**Clause**		**Free adjunct**	
	#	%	#	%	#	%	#	%	#	%
kind of	6	33.3	10	55.6	1	5.6	1	5.6	0	0
sort of	15	42.9	9	25.7	1	2.9	0	0	10	28.6
Total	21	**39.6**	19	35.9	2	3.8	1	1.9	10	18.9

As seen from Table 4, the type nouns modified verbs most frequently, adjectives taking the second place. The findings are in line with Aijmer (2002). Her results demonstrated that, for example, *sort of* "was especially frequent before NP (42.1 %) and before VP (36 %)" (Aijmer 2002, 183). The cases were the type noun featured as a stand-alone marker exclusively appeared with *sort of* only, which might imply that "<...> *kind of* strongly prefers to modify nouns and adjectives while *sort of* strongly prefers to modify whole propositions, adverbs and verbs" (Gries, David 2007).

The investigation of *kind of* and *sort of* in the EN–trans texts only confirmed the tendency. Even fewer cases of the non–internal use of the type nouns were identified – out of 200 cases (100 of *kind of* and 100 of *sort of*) there were only three cases of non–internal use: one case where *sort of* modified an adjective, one case where *kind of* modified a verb and one case where it modified an adjective, which makes only 1.5 % of the NP non–internal use. The rest 88.5 % of the concordance witnesses NP–internal use.

4.1.2. The NP–internal use of *kind of* and *sort of*

As has already been mentioned, the prevailing use of *kind of* and *sort of* in the EN–orig and EN–trans texts was NP–internal. Table 5 and Table 6 present the distribution of the NP–internal use in accordance with the propositional use indicating 'a type of' and the qualifying use when the items under study do not show any referential potential but function as discourse markers, downtoners or hedges (Quirk et al. 1985; Biber et al. 1999; Keizer 2007; Dehé, Stathi 2016).

Table 5. The frequency of the NP–internal *kind of* vs. *sort of* in the EN–orig texts

	Propositional 'type of'		DMs		Total #
	#	%	#	%	
kind of	115	63	67	37	182
sort of	63	38	102	62	165
Total #	148		201		347[5]

---

[5] The total number 347 shows the NP–internal use; 53 cases included the NP–external use.

Table 6. The frequency of the NP-internal *kind of* vs. *sort of* in the EN-trans texts

	Propositional 'type of'		DMs		Total
	#	%	#	%	#
kind of	67	68 %	31	32%	98
sort of	39	40%	60	60%	99
Total #	106		91		197[6]

As is obvious from the tables above, *sort of* conveyed pragmatic meanings more often than *kind of* in both sub-corpora. Its use as a discourse marker is almost twice more frequent than that *of kind* of (62 % vs. 37 % in the EN-orig texts and 60 % vs. 32 % in the EN-trans texts). The findings are in line with Aijmer (2002), who noted that "*kind of* is not used to the same extent as *sort of* at least on the British scene" (Aijmer 2002, 207). In her data *kind of* was used as a clear discourse marker only in 17.8 %. The reason for this is hard to find. One of the tentative observations might be the findings presented by Brems and Davidse (2010, 193), whose diachronic analysis shows a constant growth of the use of *kind of* in the binominal construction from the year of 1780 to the present day.

4.2. Translational correspondences and functions

4.2.1. The NP-internal *kind of* and *sort of* as 'a type of'

The whole range of TCs of *kind of* and *sort of* denoting a type of something is presented in Table 7. The TCs *kažkoks*, *kažkokia*, *kažkokie*, *kažkokių* 'somewhat' or, for example, *tie*, *tų*, *toms* 'these' and other forms were regarded as a single form as Lithuanian is an inflected language showing case, number and gender agreement. Since there seemed to be no variation in the TCs of *kind of* and *sort of* in the two sub-corpora, all of them are summed up together.

---

[6] The total number 197 shows the the NP-internal use; 3 cases included the NP-external use.

Table 7. TCs of *kind of* and *sort of* as 'a type of' in both sub-corpora

TCs	#	%
toks/šitoks/toks pat 'this/that/these/those/the same'	101	39.8
koks 'what kind of'	57	24
tas/tie/šie 'this/that/these/those'	17	7
kas 'what/who'	5	2
kažkoks 'some'	5	2
kitoks 'different'	4	1.6
rūšis/tipas/kategorija 'kind/type/category'	3	1.2
įvairiausios 'various'	1	0.4
savotiškas 'peculiar'	1	0.4
ø	60	24
Total	254	100

Table 7 demonstrates that the congruent TCs of *kind of* and *sort of* indicating a type are very rare (1.2 %). There are only 3 cases where they are rendered by *rūšis* 'species', *kategorija* 'category', and *tipas* 'type' (see (19)):

(19)   EN-orig: *He was **that kind of guy**.*

LT-trans: ***Tokiai vaikinų kategorijai** jis priklausė.*

The majority of translational correspondences were divergent and included demonstrative pronouns such as *tie/šie* 'those/these' and *toks/šitoks* 'such', which are generally used to refer back to some information mentioned before or something that will be presented later:

(20)   LT-orig: *Blogiausia, kad **tokios** neapykantos neišperka meilė.*

EN-trans: *The worst of it is that love doesn't compensate for that **kind of** hatred.*

(21)   EN-orig: *I always thought that a veruca was **a sort of wart** that you got on the sole of your foot!*

LT-trans: *Aš visada maniau, kad Veruka - **tai tokia karpa,** kuri atsiranda ant kojos pado!*

The results show that wh-questions with *kind of* and *sort of* quite frequently denoted a type and their Lithuanian TCs mostly encompass *koks* or *kas* 'what/who', for example:

(22)   LT-orig: ***Koks** gali būti nuolankumas šiais laikais, kai vyrų ir moterų teisės lygios!*

EN-trans: *What **kind of** humility can there be these days when men and women have equal rights!*

The findings seem to coincide with the ones obtained by Janebová and Martinková (2017), as their study unveiled that in such contexts the Czech equivalents of the type nouns "were phoric expressions (demonstrative pronouns, phoric expressions meaning "similar" or "same", and adjectives and adverbs referring to the common ground such as "typical", "precise/ly", or "definite"" (Janebová, Martinková 2017, 205).

4.2.2. *Kind of* and *sort of* as discourse markers

The TCs of *kind of* and *sort of* functioning as discourse markers in the EN-orig texts are displayed in Table 8 below. The cases of their non-internal use (i.e. the adverbial constructions) have been added to the total.

Table 8. TCs of *kind of* and *sort of* as DMs in the LT-trans texts (EN-orig → LT-trans)

TCs	#	%
kažkas/kažkoks/koks nors 'something like'	51	20
tartum/tarsi/lyg (ir)/it 'as if'	37	15
savotiškas 'peculiar'	20	8
panašus/panašiai 'similar(ly)'	6	2
maždaug taip/mažumėlę 'approximately'	4	1.6
kažkaip 'somewhat'	4	1.6
galima sakyti/pavadinti 'it can be said'	3	1.2
kaip ir 'somehow'	2	0.8
beveik 'almost'	2	0.8
tikra 'real'	1	0.4
visai 'totally'	1	0.4
atrodo 'it seems'	1	0.4
gana 'rather'	1	0.4
ø	121	48
Total	254	100

A great variety of TCs reveal the multifunctional nature of the two items in question. Naturally, the translational correspondences of the two DMs could not be congruent. The divergent TCs included a variety of words, such as *kažkoks/koks nors* 'something like' (20 %), *tartum/tarsi*, *lyg (ir)* and *it* 'as if' (15 %), *savotiškas* 'peculiar' (8 %). The other various TCs make up (9.6 %).

The translational profile of *kind of* and *sort of* as DMs in the EN-trans texts yielded somewhat similar results (see Table 9); however, the range of TCs is much narrower. The three cases of their non-internal use (i.e. the adverbial constructions) have been added to the total, too. The reason for a less divergent translational profile might be the fact that Lithuanian does not have fully grammaticalized equivalents for the English type nouns; other means of expression of approximation or mitigation may be in use, which did not fall under the scope of the paper.

Table 9. TCs of *kind of* and *sort of* as DMs in the LT-orig texts (EN-trans → LT-orig)

TCs	#	%
kažkas/kažkoks 'something like'	20	21
kažin koks/kokia 'somewhat'	17	18
šioks toks/tokia 'this/that'	9	9.6
tarsi/lyg (ir)/it 'as if'	6	6
savotiškas 'peculiar'	3	3
ø	39	41
Total	94	100

As seen from the tables, in both sub-corpora the pronominal strategy of TCs predominates. The most frequent TCs remain the same: *kažkas/kažkoks* 'something like' (20 % and 21 %). The great diversity of TCs of *kind of* and *sort of* might suggest that there is no closely related one-to-one correspondence between the two DMs in question and their Lithuanian equivalents as is the case of the DM *well* and its prototypical Lithuanian counterpart *na* (Šolienė 2020). Her study showed that almost half (47.3 %) of the cases of the use of *na* were rendered as *well* into English. The strikingly high percentage of zero correspondence (48 % and 41 %) is also indicative of the functional versatility as well as of the absence of direct equivalents of *kind of* and *sort of* in Lithuanian. The phenomenon of zero correspondence was also attested for *kind of* and *sort of* in their propositional meaning when they denoted a type (24 %). It will be separately discussed in sub-section 4.2.3 below.

When it comes to the correlation of the function of *kind of* and *sort of* in discourse and their Lithuanian TCs, a number of observations can be made. First of all, the two items under study can perform textual and interpersonal functions. For example, Aijmer (2002, 191) claims that *sort of* can be used epistemically (evidentially) to mark imprecision and self repair and affectively (interpersonally) to downtone and hedge strong opinions. It must be noted that a clear boundary between different functions is hard to draw since very often they overlap (Beeching 2016, 158–159).

The pronominal strategy in translation may be so frequent because pronouns can perform a wide range of functions, especially when they are combined with the particles *lyg* or *tarsi*. For instance, *koks, kažkoks*, *kažin koks* ‘something like‘ allow the speaker to indicate that the concept is in some way peculiar and he/she cannot recognize it or tell exactly what it is:

(23)   EN-orig: *Professor Lupin had compiled the most unusual exam any of them had ever taken; **a sort of obstacle course** outside in the sun, where they had to wade across a deep paddling pool <...>.*

LT-trans: *Profesorius Lubinas jį parengė neregėtai įdomiai: kieme, saulės atokaitoje, jų laukė **tarsi koks kliūčių ruožas**. Jie turėjo perbristi gilų baseinėlį <...>.*

(24)   EN-orig: *<...> I didn‘t feel terror so much as **a kind of vague queasiness**.*

LT-trans: *<...> siaubo nejaučiau, tik **kažkokį miglotą nerimą**.*

In (23) and (24) the speaker is looking for the appropriate word to describe a range of obstacles or to name a feeling at the same time conveying his/her subjective attitude.

The adjectival TCs of *kind of* and *sort of* show that the speaker keeps the description at a fuzzy level, too. The most frequent TCs are *savotiškas* ‘peculiar‘ and *panašus* ‘similar‘, which express non-typicality, vagueness and imprecision:

(25)   EN-orig: *He thought with **a kind of astonishment** of the biological uselessness of pain and fear <...>.*

LT-trans: *Jis su **savotiška nuostaba** pagalvojo apie biologinį skausmo ir baimės nereikalingumą <...>.*

(26)   EN-orig: *He remembered a cell with a plank bed, **a sort of shelf** sticking out from the wall, and a tin wash-basin <...>.*

LT-trans: *Jis prisiminė kitą kamerą su lentine lova, **panašia į lentyną** prie sienos, ir skardine praustuve <...>.*

Interestingly, the analysis revealed that the phrase *kind/sort of thing* was quite common in the corpus. Beeching (2016, 158) calls this lexicalised chunk as "the general extender" and maintains that it "also serves, retroactively, to hedge or downplay any perceived conceitedness and thus serves as a face-threat mitigator". In my dataset it usually appeared as part of a sentence (27) or as a stand-alone unit (28). The TCs included phrases such as *panašiai* 'similarly', *panašūs dalykai* 'similar things' *kažkas tokio* 'something like' and *maždaug taip* 'approximately', which is also a means to convey vagueness and imprecision:

(27)   EN-orig: *Not in a bad way, just to calm someone down, **that kind of thing**.*

LT-trans: *Ne blogąja prasme, tik norėdamas nuraminti kitus, **kažkas tokio**.*

(28)   EN-orig: *Black eyes. Missing teeth. **That sort of thing**.*

LT-trans: *Su mėlynėmis po akims. Be dantų. **Ir panašiai**.*

The greatest variety of TCs was attested in the intersubjective domain: *tartum/tarsi/lyg (ir)/it* 'as if', *kažkaip* 'somewhat', *galima sakyti/pavadinti* 'it can be said', *kaip ir* 'somehow', *atrodo* 'it seems', etc. The most frequent TCs were *tartum/tarsi/lyg (ir)/it* 'as if' which seem to be the closest equivalent to *kind of/sort of* as they may equally show inexactness; serve as discourse markers that weaken the illocutionary force and flag the speaker's uncertainty, as in (29)–(31):

(29)   EN-orig: *"He never really got to the point."*

*"You looked **kind of** mad," she fished.*

"Did I?" I kept my expression blank.

LT–trans: *– Jis taip tiesiai ir neprisipažino.*

*– Atrodei **lyg ir** supykusi, – tardė ji.*

– Tikrai? – nustebusi pažiūrėjau į ją.

(30)   LT-orig: *Turi **lyg ir** savo savivaldą.*

EN-trans: *It has its own **sort of** self-rule.*

(31)   EN-orig: *My great-grandpa, Ephraim Black, was **sort of** the last chief we had <...>.*

LT-trans: *Mano senelis Efraimas Blekas buvo **tarsi** paskutinis mūsų vadas <...>.*

In the examples above the speaker tries to make his/her opinion sound safely vague and thus softens the effect of his/her assertion. Example (29) is also a case of the speaker's unwillingness to impose on the hearer, by using *lyg ir* 'as if' he/she tries to save the interlocutor's face.

Finally, there are TCs such as *beveik* 'almost', *gana* 'somewhat', and *kažkaip* 'somehow', which downtone the proposition, but at the same time explicate the meaning of imprecision and approximation, for example:

(32)   EN-orig: *I've moved out. **Sort of**, anyway.*

LT-trans: *Na, **beveik** išsikrausčiau.*

(33)   EN-orig: *It was **kind of** childish, really.*

LT-trans: *Tai buvo **gana** vaikiška.*

The two units in question may also signal a close relationship between the speakers. They "make inferences or assumptions about each other's knowledge" (Aijmer 2002, 202), for example:

(34)   EN-orig: *"He works at Hogwarts."*

*"Oh," said the boy, "I've heard of him. He's **a sort of** servant, isn't he?"*

"He's the gamekeeper," said Harry.

LT-trans: – *Dirba Hogvartse.*

– *Aaa, girdėjau. Jis ten **lyg ir** tarnas?*

 – *Sargas, – atšovė Haris, vis labiau nemėgdamas vaikėzo.*

In example (34) the boy makes an inference that a person is a servant based on the hearsay and general knowledge and at the same time seeks a confirmation from his interlocutor. The Lithuanian correspondence *lyg ir* 'as if' mitigates the question. In such contexts the Lithuanian correspondences also include *galima sakyti* 'it can be said':

(35)   EN-orig: *"If we are dealing with newborns, he'll be helpful."*

"Jasper? Why?" Edward smiled darkly.

*"Jasper is **sort of** an expert on young vampires."*

LT-trans: – *Jis padės susitvarkyti su naujagimiais.*

– *Džasperis? Kaip? – Edvardas niūriai šyptelėjo.*

– *Džasperis, **galima sakyti**, yra jaunų vampyrų specialistas.*

Even though Edward does not seem to have a clue why Jasper may be useful in dealing with newborns, the initiator of the dialogue appeals to common knowledge that he wants to share. The implication is that in general Jasper is known to be a good specialist when it comes to

dealing with young vampires. Here *galima sakyti* 'it can be said' could be substituted by *žinok* 'you know'.

The varied list of TCs may prove that Lithuanian does not have a single equivalent for the two multifunctional units discussed herein. Also, it is sometimes difficult to disentangle different functions performed by *kind of* and *sort of* since they overlap (Beeching 2016, 160).

4.2.3. Zero correspondence

Overall, zero correspondence is a unifying feature of DMs in cross-linguistic studies (Johansson 2007, 26). Due to their extreme multifunctionality, non-propositionality, context-dependence and non-referential (textual and interpersonal) function, they exhibit a wide array of different TCs and their exact cross-linguistic equivalent is hard to find. The cases of zero correspondence of *kind of* and *sort of* as DMs amount to 48 % in the LT-trans texts. Even though they had no particular source in the LT-orig texts, the two DMs were also inserted in the English translations quite frequently (41 %). The propositional type noun usage in the present study also showed a relatively high percentage of zero correspondence (24 %). Similar findings are presented in Janebová and Martinková (2017, 188): even more than a half of *kind of* and *sort of* uses have no Czech correspondences in their dataset.

One of the factors determining zero correspondence of the type nouns might be the redundancy factor and the language norm, e.g.:

(36)   EN-orig: *His lips pulled back over his teeth and his eyes shone with an odd light – a wild, fierce **kind of hope***.

LT-trans: *Dantys prasišiepė, akyse blykstelėjo keista ugnelė – nuožmi, beprotiška ø viltis.*

Example (36) is typical of the pattern *modifier + kind of/sort of + NP*, which is especially inclined to have *kind of/sort of* omitted in translation, seemingly because Lithuanian does not require an additional mediator between the modifiers and the modified head.

In addition, the language norm might also be an explanation why *kind of* and *sort of* are inserted into the English translation; it seems that it is more natural to resort to mitigation strategies in English:

(37)   LT-orig: – *øDurnystė visa tai, – sako ilgšis, – verčiau parodom jam jo nupjautą galvą.*

EN-trans: *"This is **some kind of idiocy**," says the beanpole. "It'd be better if we showed him his chopped-off head."*

It is important to note that sometimes the meaning conveyed by *kind* of and *sort of* is completely reversed in translation, i.e. a mitigated English statement is rendered as an assertion into Lithuanian, for instance:

(38)   EN-orig: *It's **kind of obvious**, when you think about it <...>.*

LT-trans: *Geriau pagalvojus, dalykas **visiškai aiškus** <...>.*

(39)   EN-orig: *Potions lessons were turning into a **sort of weekly torture**, Snape was so horrible to Harry.*

LT-trans: *Nuodų ir vaistų pamokos virto **tikra kankyne** – Sneipas ėste ėdė Harį.*

Kind of in (38) is used to mitigate face threats. By opting for *kind of obvious* the speaker has an intention to save his/her or the interlocutor's face. *Kind of* here softens a strongly voiced opinion, whereas, on the contrary, *visiškai* 'completely' shows sheer confidence and marks a

reinforced statement. In (39) *tikra* 'real' functions as an intensifier; it exaggerates the semantics of the already emotionally charged word *kankynė* 'torture'.

5. Conclusion

The present paper reports on the multifunctional English type nouns *kind of* and *sort of* and their Lithuanian correspondences in a contrastive perspective. There have been a lot of monolingual studies that analysed the English type nouns in terms of their structure or function in different perspectives, but contrastive corpus-based studies comparing two or more languages are rather scarce. This paper aims to describe the quantitative and qualitative distribution of the English *kind of* and *sort of*, to determine their translational correspondences in Lithuanian, as well as to reveal their functional diversity. The research method is a quantitative and qualitative contrastive analysis based on data extracted from the self-compiled bidirectional corpus ParaCorp$_{EN \to LT \to EN}$ comprising fiction texts.

The results show *kind of* and *sort of* are prone to be used NP-internally; however, even in this construction they can feature as DMs. *Kind of* and *sort of* function as unambiguous DMs when they completely lose their nominality, i.e. are used NP-externally. *Sort of* conveyed pragmatic meanings more often than *kind of* in both sub-corpora, whereas *kind of* was more frequently used propositionally to indicate a type.

The functional and semantic potential of the type nouns is fully reflected by their TCs. Very rarely *kind of* and *sort of* are translated congruently into a type noun (*rūšis* 'kind', *kategorija* 'category' and *tipas* 'type'); even though they denote a type and are used propositionally, they usually correspond to demonstrative pronouns such as *tie/šie* 'those/these' and *toks/šitoks* 'such' in Lithuanian.

The translational profile has revealed that certain functions of *kind of* and *sort of* as DMs, such as textual and interpersonal, may be realised by different Lithuanian TCs which may help establish the common ground between the speaker and the hearer or refer to the previous

context, may indicate epistemic imprecision, approximation or may downtone a proposition (cf. Janebová, Martinková 2017, 209). The most common TCs included *kažkas/kažkoks* 'something like', *kažin koks/kokia* 'somewhat', *tartum/tarsi/lyg (ir)/it* 'as if', *beveik* 'almost', *gana* 'somewhat', *kažkaip* 'somehow', etc. Finally, the same TCs may appear in both functions, which may be indicative of the functional overlap.

The high number of zero correspondence has also revealed some differences between English and Lithuanian. First of all, it is obvious that due to the multifunctional nature and context-dependence of *kind of* and *sort of* their exact cross-linguistic equivalents are hard to find. Second, it signals that the Lithuanian type nouns have not advanced on the grammaticalization path the way the English type nouns have and there is no one-to-one correspondence of the forms. Moreover, it seems that the two languages do not put the same emphasis on the mitigation and politeness strategies. Discourse markers are a feature of vague language; their precise meaning is difficult to pin down, so they disappear in translation very frequently.

As this particular research was based on a quite limited number of fiction texts, the analysis of larger corpora encompassing different registers, for example, spoken or academic, would provide more interesting and insightful results. Further research could also include the plural forms *kinds of* and *sorts of*. Moreover, the most frequent Lithuanian TCs of *kind of* and *sort of toks/šitoks* 'such' and *kažkas/kažkoks* 'something like' should not escape further analysis.

List of Abbreviations

#	– raw frequency
DM(s)	– discourse marker(s)
EN	– English
f	– normalized frequency
LT	– Lithuanian
N2	– the second noun within the noun phrase
NP	– noun phrase
Ø	– zero correspondence
orig	– original texts
SL	– source language
TC(s)	– translational correspondence(s)
TL	– target language
trans	– translated texts
VP	– verb phrase

Data sources

ParaCorp$_{EN \rightarrow LT \rightarrow EN}$ Bidirectional Parallel Corpus of English and Lithuanian (Šolienė 2013)

Acknowledgments

critical remarks on the original draft, which have substantially improved the final version of the paper. Any errors that remain are my sole responsibility.

References

Aijmer Karin 2002, *English Discourse Particles: Evidence from a Corpus*, Amsterdam: John Benjamins.

Aijmer Karin, Simon-Vandenbergen Anne-Marie 2003, The discourse particle *well* and its equivalents in Swedish and Dutch, *Linguistics* 41(6), 1123-1161.

Barlow Michael 1995, *A Guide to ParaConc*, Athelstan: Houston.

Beeching Kate 2016, *Pragmatic Markers in British English: Meaning in Social Interaction*, Cambridge: Cambridge University Press.

Beeching Kate, Detges Ulrich (eds.) 2014, *Functions at the left and right periphery: Crosslinguistic investigations of language change*, Leiden: Brill.

Biber, Douglas, Johansson Stig, Leech Geoffrey, Conrad Susan, Finegan Edward (eds.) 1999, *Longman Grammar of Spoken and Written English*, London: Longman.

Brems Lieselotte 2011, *Layering of size and type noun constructions in English*, Berlin/Boston: Mouton de Gruyter.

Brems Lieselotte, Davidse Kristin 2010, The Grammaticalisation of Nominal Type Noun Constructions with *kind/sort of*: Chronology and Paths of Change, *English Studies* 91(2), 180-202.

Davidse Kristin, Brems Lieselotte, De Smedt Liesbeth 2008. Type noun uses in the English NP. A case of right to left layering, *International Journal of Corpus Linguistics* 13(2), 139–168.

Degand Liesbeth 2009, On describing polysemous discourse markers. What does translation add to the picture? In *From will to well. Studies in Linguistics offered to Anne-Marie Simon-Vandenbergen*, Slembrouck Stef, Taverniers Miriam, Van Herreweghe Mieke, eds., Gent: Academia Press, 173-183.

Dehé Nicole, Stathi Katerina 2016, Grammaticalization and prosody: The case of English sort/kind/type of constructions, *Language* 92(4), 911-946.

Denison David 2005, *The Grammaticalisations of Sort of, Kind of and Type of in English*. Paper presented at New Reflections on Grammaticalization 3, Santiago de Compostela, July 17–20. http://www.humanities.manchester.ac.uk/medialibrary/llc/files/david-denison/Santiago_NRG3_paper.pdf.

Dyvik Helge 2004, Translations as semantic mirrors: from parallel corpus to Wordnet, *Language and* Computers 49(1), 311-326.

Fetzer Anita 2010. Hedges in context: form and function of *sort of* and *kind of*. In *New Approaches to Hedging*, Kaltenböck Gunther, Mihatsch Witrud, Schneider Stefan, eds., Bingley: Emerald, 49-71.

Furkó Bálint Péter 2014, Perspectives on the translation of discourse markers, *Acta Universitatis Spientiae, Philologica* 6, 181-196.

Gries Stephan Th., David Caroline V. 2007, This is kind of/sort of interesting: variation in hedging in English, In *Studies in Variation, Contacts and Change in English. Volume 2. Towards Multimedia in Corpus Studies*, Päivi Pahta, Taavitsainen Irma, Nevalainen Terttu, Tyrkkö Jukka, eds., https://varieng.helsinki.fi/series/volumes/02/gries_david/.

Holmes Janet 1988, *Sort of* in New Zealand Women's and Men's Speech, *Studia Linguistica* 42(2), 85–121.

Janebová Markéta, Martinková Michaela 2017, NP-Internal *Kind of* and *Sort of*: Evidence from an English-Czech Parallel Translation Corpus, In *Contrasting English and Other Languages through Corpora*, Janebová Markéta, Lapshinova-Koltunski Ekaterina, Martinková Michaela, eds., Newcastle upon Tyne: Cambridge Scholars Publishing, 164-218.

Johansson Stig 2007, *Seeing through Multilingual Corpora: On the use of corpora in contrastive studies*, Amsterdam/Philadelphia: John Benjamins Publishing Company.

Kay Paul 1984, The *Kind of/Sort of* Construction, In *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*, 157–171.

Keizer Evelien 2007. *The English Noun Phrase: The Nature of Linguistic Characterization*, Cambridge: Cambridge University Press.

Kirk John M. 2015, *Kind of* and *sort of*: Pragmatic discourse markers in the SPICE-Ireland Corpus, In *Pragmatic Markers in Irish English*, Amador-Moreno Carolina P., McCafferty Kevin, Vaughan Elaine, eds., Amsterdam/Philadelphia: John Benjamins Publishing Company, 89-113.

Kolyaseva Alena, Davidse Kristin 2016, A typology of lexical and grammaticalized uses of Russian *tip*, *Leuven Working Papers in Linguistics* (5), 171–210.

Lewis Diana 2006, Contrastive analysis of adversative relational markers, using comparable corpora, In *Pragmatic Markers in Contrast*, Aijmer Karin, Simon-Vandenbergen Anne-Marie, eds., Oxford: Elsevier, 139–153.

Margerie Hélène 2010, On the rise of (inter)subjective meaning in the grammaticalization of *kind of /kinda*, In *Subjectification, intersubjectification and grammaticalization*, Davidse Kristin, Vandelanotte Lieven, Cuyckens Hubert, eds., Berlin: Mouton de Gruyter, 315–346.

Quirk Rudolf, Greenbaum Sidney, Leech Geoffrey, Svarvik Jan 1985, *A Comprehensive Grammar of the English Language*, London: Longman.

Šolienė Audronė 2013, *Episteminio modalumo ekvivalentiškumo parametrai anglų ir lietuvių kalbose*. Humanitarinių mokslų daktaro disertacija, Vilnius: Vilniaus universitetas.

Šolienė Audronė 2020, Lithuanian discourse markers *na* and *nu*: A glimpse at Lithuanian-English parallel corpus data, In *Translating and Comparing Languages: Corpus-based Insights*, Granger Sylviane, Lefer Marie-Aude, eds., Louvain-la-Neuve: Presses universitaires de Louvain, 237–255.

Traugott Elizabeth Closs 1995, Subjectification in grammaticalisation, In *Subjectivity and Subjectivisation: Linguistic Perspectives*, Stein Dieter, Wright Susan, eds., Cambridge: Cambridge University Press, 31–54.

Traugott Elizabeth Closs 2008, Grammaticalization, constructions and the incremental development of language: Suggestions from the development of degree modifiers in English, In *Language Evolution: Cognitive and Cultural Factors*, Eckardt Regine, Jaeger Gerhart, Veenstra Tonjes, eds., Berlin: Mouton de Gruyter, 219-240.

Traugott Elizabeth Closs 2010, (Inter)subjectivity and (inter)subjectification: A reassessment, In *Subjectification, Intersubjectification and Grammaticalization*, Davidse Kristin, Vandelanotte Lieven, Cuyckens Hubert, eds., Berlin: De Gruyter, 29–71.

Usonienė Aurelija 2006, Episteminio modalumo raiška. Ekvivalentiškumo bruožai anglų ir lietuvių kalbose, *Darbai ir dienos* 45, 97–108.

Wichmann Anne, Simon-Vandenbergen Anne-Marie, Aijmer Karin 2010, How prosody reflects semantic change: A synchronic case study of *of course*. In *Subjectification, intersubjectification and grammaticalization*, Davidse Kristin, Vandelanotte Lieven, Cuyckens Hubert, eds., Berlin: Mouton de Gruyter, 103–54.