Juk and *gi*, and "particles" in contemporary Lithuanian: Explaining language-particular elements in a cross-linguistic context[1]

Vladimir Panov

General Linguistics Department, Faculty of Philology
Vilnius University
Universiteto g. 5
LT-01513 Vilnius, Lithuania
E-mail: vladimir.panov@flf.vu.lt

Institute of Linguistics
Russian Academy of Sciences
1 bld. 1 Bolshoi Kislovsky lane
125009 Moscow, Russian Federation

Abstract. In this paper, I propose to take a fresh look at the elements traditionally termed "particles" in descriptions of Lithuanian. I establish criteria that would allow one to come up with a meaningful classification of such elements, and argue for the descriptive language-particular category of "sentence particles". I then turn to the syntax and functions of *juk* and *gi*, two prominent elements of this category. In order to understand their functions better, I invoke cross-linguistic comparison. I demonstrate that the ideas expressed in the literature on the German particles *doch* and *ja* can also be applied to Lithuanian. All four particles exhibit a common meaning component, which I describe as *framing the proposition as uncontroversial*, but also significant differences. Finally, I provide a brief overview of similar categories in some neighboring and genealogically related languages, and introduce the general methodological and areal-typological dimension of the phenomenon.

Keywords: discourse particles, uncontroversial information marking, shared knowledge, common ground, illocutionary force, Lithuanian, German, comparative concepts

1 Introduction

1.1 Preliminaries

In both descriptive and comparative linguistics, the study of the level of discourse and "pragmatics" has long been on the periphery. Multiple grammatical descriptions of particular languages containing fine-grained analysis of morphology and syntax are sometimes strikingly incomplete in the parts devoted to what is usually labeled as "particles", "pragmatic markers", "discourse particles", "discourse markers", etc.

In typology, the advanced cross-linguistic study of discourse-associated elements can be traced back to the 1980s, an important milestone being the edited volume by Weidt (1989). Not surprisingly, the studies of discourse particles etc. evolved in close association with the linguistics of German: the extensive use of so-called "modal particles" (*Modalpartikeln, Abtönungspartikeln*) is a hallmark of this language.

The objective of this paper is twofold. First, I provide a working typologically-informed framework for analyzing the syntax and functions of the elements of Lithuanian which are normally termed "particles" in the literature. Within this framework, I focus on two "particles" of contemporary Lithuanian, namely *juk* and *gi*, whose functions and morphosyntactic[2] properties have never been given sufficient attention. *Juk* and *gi* are considered together, as they exhibit a certain degree of functional overlap, although there are also important differences. Giving an account of the functioning of *juk* and *gi* is my second and main goal. In doing this, I invoke cross-linguistic comparison, namely, the studies of German particles *doch* and *ja*. The latter exhibit interesting functional parallels with *juk* and *gi* but are much better described in the literature.

Although this study is primarily descriptive, it also appeals to the typological community, and addresses the issue of commensurability between language-particular structures. The reader can find some discussion of the relevant methodological debate in 2.1 and 4.

1.2 Structure of the paper

In Section 2.1, I address the issue of terminology in the studies of discourse-associated elements. Section 2.2 discusses the morphosyntactic classes of "particles" in Lithuanian, and a new classification of this kind of elements is proposed. In 2.3, I take a closer look at the positioning of *juk* and *gi*. Section 3.1 discusses the morphosyntax and the functions

---

[2] In this paper, I use the term "morphosyntactic" in the sense "morphological and/or syntactic", following Haspelmath (2011). Both cross-linguistically and language-specifically, it is often very difficult if possible at all to draw a clear line between "morphological" and "syntactic", which are better seen as a "morphosyntactic" continuum.

of the particles *doch* and *ja* in German. In 3.2, I analyze the functional contribution of *juk* and *gi* in Lithuanian. In 3.3, the particles of German are compared with the Lithuanian ones. In 4.1, I highlight the contribution of the present study to the current methodological debate on "comparative concepts" (Haspelmath 2010) in linguistic typology. In 4.2, the areal-typological and diachronic dimension of the functions exhibited by *juk* and *gi* is briefly introduced. In the Conclusion, I summarize the findings of the study.

2 *Juk*, *gi*, and the "particles" of Lithuanian: Morphosyntax

2.1 Terminological issues

In the last decade, there has been some debate on the terms used for markers exhibiting functions usually ascribed to the domain of discourse, both cross-linguistically and language-specifically.

Both aspects of the problem – the domain of discourse itself and the labeling of markers associated with it – are problematic. First, it is hardly possible to draw a clear boundary between "discourse" and "non-discourse" (e.g. grammatical) functions. In this paper, I take an agnostic position and do not attribute the functions studied here to "discourse" or any other domain. When describing the functioning of *juk* and *gi* in Lithuanian, this would not contribute any additional information.

The most popular labels for the markers associated with the domain of discourse, regardless of what authors have in mind, are "discourse markers", "modal particles", "discourse particles", "pragmatic particles", and many others. There are numerous works which discuss the differences between these labels, the most influential recent cross-linguistic studies (in non-generative linguistics) being Degand et al. (2013) and Fedriani & Sansò (2018). However, the problem with both volumes as well as many other works is the aprioristic approach, which has been extensively criticized by Haspelmath (2007, 2010, 2018). Although this stance is not always made explicit, authors believe that there are natural-kind-like cross-linguistic categories such as "discourse markers" or "modal particles" that manifest themselves in various ways in different languages, and one can "diagnose" them attributing language-particular elements to one or another cross-linguistic kind. For this paper, I refuse such an approach and take a non-aprioristic stance in this respect, accepting the view that morphosyntactic criteria are language-particular.

This does not mean, however, that the categories of various languages are completely different and incommensurable. On the contrary, the morphosyntactic categories of different languages sometimes are strikingly similar, especially when the languages are genealogically related or geographically adjacent. In such cases, it is indeed convenient to use a single term in order to talk about the categories of different languages, being aware that there are, however, differences between them, and highlighting these differences.

2.2 What are "particles" in Lithuanian?

As many other language-particular descriptions, the most widely used comprehensive grammar of Lithuanian (Ambrazas 1997) contains a section entitled "particles" (Lith. *dalelytės*). This section is an overview of the elements with a very broad scale of morphosyntactic and semantic properties. In fact, what unites them is that these elements do not fall easily into other function-word classes such as conjunctions, connectives, or propositions. These elements are typically short (1-2 syllables). The criteria for semantic and syntactic classification of particles provided in the grammar are not made explicit, and sometimes the motives for grouping elements in one or another way remain obscure.

The elements we focus on in this paper are classified as "intensifying-emphatic particles" in Ambrazas (1997, 401–402). Besides *juk* and *gi*, the connecting element *ir* is also included in this group. All three "emphasize a word or a clause". In my view, there are at least two problems with this definition. First, it is not made explicit what is understood under "emphasis": this notion seems far too vague and permits a large range of interpretations. Second, modifying a "word" or a "clause" are two very different things in terms of semantic scope. In fact, in the examples of *gi* in which it is claimed to intensify a single word, the particle's scope is the whole sentence ("The particle *gi*, usually postposed *to the word it intensifies*, sometimes occurs in the initial position" [Ambrazas 1997, 402]):

(1)   a.[3] *Žinai*   ***gi***   *jo*   *papročius.*[4]
          know.PRS.2   **gi**   he.GEN   ways.ACC.PL

          'You know his ways.'

    b. *Gi* *žinai,*   *ko*   *jam*   *reikia.*
         gi   know.PRS.2   what he.DAT is.needed.PRS.3

         'You do know what he wants.' (Ambrazas 1997, 402)

In both examples, the particle occurs in direct contact with the verb following (1a) or preceding it (1b). In (1b), the first clause consists of a single verb predicate. If the scope of *gi* were the "word", or a single constituent consisting of a single head word which is a verb in this case, one would suspect *gi* to be able to take another constituent as its scope if it is moved to a different position. If we take a slightly longer sentence, one can easily observe that this is not the case.

---

[3] I this paper, I adopt Salos Glossing Rules (Nau & Arkadiev 2015).
[4] Lithuanian and German examples of the present paper for which no sources are indicated were elicited from Lithuanian native speakers (1 female and 2 males around 30 years old, and 1 female around 60). The informants generally agreed in their intuitions. In the rest of the cases, the sources are explicitly given.

(2)   a. *Tu*   *gi*   *mėgsti*   *valgyti*   *saldžius*
          2SG.NOM   **gi**   like.PRS.2   eat.INF   sweet.ACC.PL[M]

          pyragaičius   *ir*   *sultingas*   *braškes.*
          pastry.ACC.PL   and   juicy.ACC.PL[F]   strawberry.ACC.PL

      b. *Tu*   *mėgsti*   *gi*   *valgyti*   *saldžius*
          2SG.NOM   like.PRS.2   **gi**   eat.INF   sweet.ACC.PL[M]

          pyragaičius   *ir*   *sultingas*   *braškes.*
          pastry.ACC.PL   and   juicy.ACC.PL[F]   strawberry.ACC.PL

      c. *?Tu*   *mėgsti*   *valgyti*   *saldžius*
          2SG.NOM   like.PRS.2   eat.INF   sweet.ACC.PL[M]

          pyragaičius   *ir*   *sultingas*   *braškes*   *gi*
          pastry.ACC.PL   and   juicy.ACC.PL[F]   strawberry.ACC.PL   **gi**

      d. *??Tu*   *mėgsti*   *valgyti*   *gi*   *saldžius*
          2SG.NOM   like.PRS.2   eat.INF   **gi**   sweet.ACC.PL[M]

          pyragaičius   *ir*   *sultingas*   *braškes.*
          pastry.ACC.PL   and   juicy.ACC.PL[F]   strawberry.ACC.PL

      e. **Tu*   *mėgsti*   *valgyti*   *saldžius*
          2SG.NOM   like.PRS.2   eat.INF   sweet.ACC.PL[M]

          pyragaičius   *gi*   *ir*   *sultingas*   *braškes.*
          pastry.ACC.PL   **gi**   and   juicy.ACC.PL[F]   strawberry.ACC.PL

          '[It's clear that] you like to eat sweet pastries and juicy strawberries.'

Ex. (2a) is the most natural and strictly preferred in contemporary Lithuanian, (2b) and (2c) are reported possible in spontaneous speech (2b is better than 2c), (2d) is highly unlikely (possible only in a self-correction), and (2e), in which *gi* follows the first direct object, is impossible. In (2c) *gi*, although it follows the second direct object, this constituent is not focused: it is still the whole sentence that the particle takes scope over, as in (2a) and (2b). One can conclude that the scope of *gi* is the whole sentence regardless of its position,[5] which is most often the second one ("Wackernagel"), but sometimes also the sentence-initial (1b), the postverbal (2b), or the sentence-final one (2c).[6]

---

[5] The situation was different in Old Lithuanian, see the Section 4.2.

[6] An anonymous reviewer argues that one could reformulate the sentence in (1a) so that *gi* follows the nominal constituent: *[[jo paproćius gi] žinai]*. However, even in this example there is no reason to believe that *gi* "stresses" *jo paproćius* and not the whole sentence. The particle is found in the second position after the first consitutent, and its scope is the whole sentence. However, *jo paproćius* is indeed focalized or topicalized here (depending on the intonation), but the information structure is markerd by fronting the constituent, not by *gi*.

Turning back to (1b), we are dealing with a biclausal construction in which the main clause consists of a single verb. Here, *gi* modifies the whole construction. This can be observed if one, e.g., extends the first clause (3) or moves *gi* to the end (4).

(3) *Tu*   *gi žinai*   *labai gerai, ko*   *jam*
     2SG.NOM   **gi** know.PRS.2   very   well   what.GEN   he.DAT

     reikia.
     be.needed.PRS.3

     'You do know very well what he wants.'

(4) *Žinai,*   *ko*   *jam*   *reikia*   *gi*.
     know.PRS.2   what.GEN   he.DAT   be.needed.PRS.3   **gi**

     'You do know what he needs.'

The same observation holds for *juk*, the main difference being the preferred position which is sentence-initial in the latter case.

Unlike *gi* and *juk*, the elements such as *ir* 'also' and *net* 'even' (included in the same "intensifying-emphatic" category in Ambrazas, 2006) exhibit a different scope-related behavior. For instance, the example with *ir* taken from the grammar can be modified in the way shown in (5):

(5) a. *Juk*   *[ir*   *[aš]]*   *tavo   duktė.*
        juk   **ir**   I.NOM   your   daughter.NOM.SG

        'I'm your daughter too [not only someone else who is not named in the sentence (reminding)]' (Ambrazas 2006, 402)

     b. *Juk*   *aš*   *[ir*   *[tavo]]*   *duktė.*
        juk   I.NOM   **and**   your   daughter.NOM.SG

        'I am your daughter too [not only the other parent's daughter]'

From the example (5), one sees a clear difference in meaning when *ir* occurs with another noun phrase, and it is the phrase and not the sentence which is in the particle's scope. *Ir* may have scope over words and phrases of different kinds; consider the various readings of the same sentence when *ir* has different scopes (6):

(6) a. *Aš*   **ir**   *picą*   *kepti*   *moku.*
        I.NOM   and   pizza.ACC.SG   bake.INF   can.PRS.1

        'I can also bake   pizza [not only bread].'

b. *Aš*   ***ir***   *picą*   *kepti*   *moku.*
I.NOM   and   pizza.ACC.SG   bake.INF   can.PRS.1

'I can also bake pizza [not only drive a car].'

The same applies to *net* 'even', which also has a narrow scope, and the interpretation of a sentence depends on the scope of *ir* or *net*.

Concerning the analysis above, I propose a set of criteria which are relevant when describing the elements of Lithuanian traditionally termed "particles"; such criteria allow us to adequately reflect their morphosyntactic properties.

– *scope*: sentence *vs* phrase. I propose to call the Lithuanian particles with sentence scope *sentence particles*[7] regardless of their preferred position in the sentence, following in part Munaro & Poletto (2004).[8] Unlike sentence particles, *phrase particles* take scope over a single constituent within their clause.[9]

– *position*. This criterion defines the preferred or obligatory position in the sentence for sentence particles and the position with respect to the constituent under the scope for phrase particles.

– *boundness*: ability *vs* inability of an element to occur in isolation. This is Bloomfield's (1933, 160) criterion which reveals the degree of morphosyntactic autonomy of an element in a language. One can also cite Haspelmath's extended formulation: free forms can form a complete (possibly elliptical) utterance (Haspelmath 2013, 213). Thus, the "particle" elements of Lithuanian can be classified as either *bound* or *free*.

– *prosodic dependency*: Integration into the intonation contour of an utterance. This criterion applies only to sentence particles. Elements can be classified as either integrated into their host sentence's intonation or non-integrated (having their own intonation pattern).

---

[7] The reader should not forget that I am talking of a language-particular descriptive category valid only for contemporary standard and colloquial urban Lithuanian. It cannot be mechanically transferred onto other languages.

[8] The cited authors use the term "sentential particles" which I find a bit clumsy in English.

[9] In the typological literature as well as in various language-particular descriptions, such elements are often addressed as "focus particles". However, there is no agreement among linguists on the boundaries of the category of "focus". It is clear that the label "focus" does not apply well to at least some phrase particles of Lithuanian.

Table 1 presents some frequently-used elements classified according to the values of the listed parameters apart from the one of prosodic dependency.[10]

	sentence scope		phrase scope
bound	preferred position in sentence		*net* 'even', *bent* 'at least', *ir* 'and/also'
	sentence-initial	second (Wackernagel)	
	juk 'uncontroversial information', *ar* 'polar question'	*gi* 'uncontroversial information'	
non-bound	*nejaugi* 'is it really the case that…', *atseit* 'unreliable information'		*būtent* 'exactly', 'just'

Table 1. Morphosyntactic types of "particles" in Lithuanian

The list presented in Table 1 is not exhaustive and only serves as an illustration of the application of some of the listed criteria. However, my suggestion is to conduct future research on Lithuanian "particles" having in mind these classification principles, avoiding the confusion of elements having very different properties.

2.3 The positions of *juk* and *gi*

I will now focus on the morphosyntax and functioning of two elements of contemporary urban spoken Lithuanian – *juk* and *gi*. Both elements fall into the type "sentence particles" if the proposed criteria are applied. Both take scope over the whole sentence, both cannot occur in isolation (are bound), both have preferred, although different, positions in their host sentences. Another and probably the single other element of Lithuanian which shares the same set of morphosyntactic properties is the polar question marker *ar* (7b), unlike its cognate and diachronic source *ar* which is a marker of alternative (7a):

(7)   a. *Aš   užeisiu   šiandien **ar**   rytoj.*
          I.NOM   come.over.1FT   today   or   tomorrow

          'I will come over today or tomorrow.'

     b. ***Ar*** *tu   užeisi   šiandien?*
          Q   2SG.NOM   come.over.2FT today

          'Will you come over today?'

---

[10] A special computer analysis of prosody would be needed here. This is beyond my scope in the present study.

Corpus data[11] indicates that *gi* strongly prefers the second position (97 of 100 random occurrences in the colloquial subcorpus of LKT), whereas *juk* occurs in the first position in (74 of 100). The notion "second position" here is to be interpreted as *the position after the first full phrase in a sentence*.[12] For example, (8a) is preferred to (8b) which sounds awkward if acceptable at all (in the case if *mano* 'my' is not focused and highlighted by the intonation):

(8) a. *Mano   brolis   gi   gyvena   Kaune.*
         my   brother.NOM.SG   **gi**   live.PRS.3   Kaunas.LOC.SG

        '[Don't forget that] my brother lives in Kaunas.'

      b. *?Mano   **gi**   brolis   gyvena   Kaune.*
         My   **gi**   brother.NOM.SG   live.PRS.3   Kaunas.LOC.SG[13]

The rare exceptions to the second-position rule for *gi* can be divided into two groups. First, *gi* sometimes starts a sentence. Out of 100 random uses in LKT (colloquial subcorpus), it occurs 3 times in the initial position.[14] An example from the corpus is:

(9) ***Gi***   *katalikybės*   *pagrindas –*   *bendruomenė.*
     Gi   Catholicism.GEN.SG   basis.NOM.SG   community.NOM.SG

    '[Don't forget that] the basis of Catholicism is community.' (LKT)

Other cases when *gi* is seemingly found in a position other than the second one may be described using the insights gained in Andrej Zaliznjak's works on the second-position clitics in Old Russian. As Zaliznjak demonstrates (2008, 47–56), Old Russian clitics are able to "move" rightward from the second position in the cases when obligatorily sentence-beginning elements such as conjunctions or fronted constituent(s) (i.e. a topicalized or a focalized phrase) are present. In such cases, a "barrier" (Rus. *ritmiko-sintaksičeskij barjer*) is placed after these elements, and the second position is counted starting from the barrier, e.g.

---

[11] Lietuvių kalbos tekstynas http://tekstynas.vdu.lt/tekstynas/, the colloquial language subcorpus (557 822 words). Accessed on 20.05.2019. Further in the paper, the same subcorpus is used, if another subcorpus is not indicated.

[12] In some other languages having a syntactically defined Wackernagel's position, the elements of this kind follow the first prosodic word, e.g. in Ancient Greek or Old Russian. For a detailed discussion see Zimmerling (2013, 44–85).

[13] In LKT (all subcorpora, 140 921 288 words), the combination *mano gi* occurs twice, and in all these cases *mano* is an independent constituent (the agent in a passive construction), and not a dependent member of a noun phrase.

[14] In Nau & Ostrowski (2010, 26), this position is mentioned as the preferred one in contemporary Lithuanian. Corpus data demonstrate that this is not actually the case, and the first position of *gi* is rare, although acceptable.

(10) *toi*   *že*   *oseni*   ‖   *mnogo* *sę*   *zla*
        that.GEN.SG[F]   PRT   autumn.GEN.SG   much   RFL   evil.GEN.SG

        sъtvori.
        do.AOR.3SG

        'That autumn many bad things happened.' (Zaliznjak 2008, 49)

Similar forces seem to be involved in Lithuanian. *Gi* also "moves" rightward in the sentence if some specific kind of element occurs in the beginning of a sentence. It may be a conjunction or a sentence-initial discourse marker as in (11a), or a fronted topicalized or focalized constituent, e.g., an adverbial as in (11b):

(11) a. *Bet*   *ko*   ***gi***   *jam*   *reikėtų*   *klausytis?*
           But   what.GEN   **gi**   he.DAT   is.needed.3SBJ   listen.INF.RFL

           'Then, whose opinion is he supposed to listen to?' (LKT)

     b. *Dabar*   *su*   *autiku*   ***gi***   *grįžau*
           now.NOM.PL   with   bus.INS.SG   **gi**   return.PST.SG.1

           'Now, I've returned with a bus [it's clear].' (LKT)

Gi sometimes (rarely) occurs sentence-finally. This use is perceived by Lithuanian speakers as non-normative and markedly colloquial. Such cases are not reflected in LKT, even in its spoken sub-corpus:

(12) *Tavo* *brolis*   *Kaune*   *gyvena*   ***gi***
        your   brother.NOM.SG   Kaunas.LOC.SG   live.PRS.3   **gi**

        '[The thing is that] your brother lives in Kaunas'

Finally, in colloquial language *gi* is able to occur after a focused constituent preserving its sentence scope. In the following example, *Kaune* 'in Kaunas' is said with an intonation which marks the sentence focus:

(13) *Tavo* *brolis*   *Kaune*   ***gi***   *gyvena*
        your   brother.NOM.SG   Kaunas.LOC.SG   **gi**   live.PRS.3

        '[The thing is that] your brother lives in Kaunas [not Vilnius]'

Ar, which is also classified as a sentence particle here, exhibits a similar behavior: its neutral position is sentence-initial (14a), but in colloquial language it is able to occur preceding the focalized constituent if there is any (14b). In (14b), the scope of *ar* is still the whole sentence: the entire sentence preserves its question illocutionary force.

(14)   a.   ***Ar***   *tavo*   *brolis*   *Kaune*   *gyvena?*
           Q   your   brother.NOM.SG   Kaunas.LOC.SG   live.PRS.3

         'Is it Kaunas where your brother lives?'

      b.   *Tavo* *brolis*   ***ar***   *Kaune*   *gyvena?*
            your   brother.NOM.SG   **Q**   Kaunas.LOC.SG   live.PRS.3

         'Is it Kaunas where your brother lives?'

In both sentences, *Kaunas* is focalized. This is marked by the focus intonation with which *Kaune* is pronounced. The sentence turns unacceptable if one preserves the focus intonation on *Kaune* but moves *ar* to another constituent, e.g., before *gyvena* 'lives'. Such a movement is only possible if *gyvena* becomes focalized, and this should be reflected in an intonation change (the focus intonation moves to *gyvena*)[15].

The particle *juk* exhibits more positional flexibility. Although its preferred position is sentence-initial (74 of 100 random occurrences in the colloquial subcorpus of the LKT), it also frequently occurs in second position. It is also able to attach to focalized constituents, either preceding or following them.

3 Understanding the functional contribution of *juk* and *gi* in a cross-linguistic perspective

3.1 *Doch & ja* of German: Morphosyntax and functions

A reason for comparing elements of different languages is the ability of these elements to translate each other. In the absence of an accessible German-Lithuanian parallel corpus, such a work cannot yet be done systematically.[16] However, a preliminary study of existing bilingual texts shows that both *gi* and *juk* often correspond to both *doch* and *ja* in translations, cf. a fragment of the Lithuanian translation of Kafka's *Die Verwandlung* (15-16):

(15)   DE: *Sie*   *verstand*   ***doch***   *alles*   *viel*
            she.NOM   understand.PST.3   **doch**   everything.ACC   much

           besser   *als*   *die*   *Schwester.*
           better   than   DEF.NOM.SG[F]   sister.NOM.SG

---

[15] Similar patterns are attested for the question particles of various languages. A well-described example is Turkish. In this language, the question particle occurs after the predicate (sentence-finally) in neutral contexts, but follows the focalized constituent in the sentences that have one (Kornfilt 1997, 438).

[16] For studies of discourse/modal particles in parallel corpora, see Aijmer and Simon-Vandenbergen (2006) and Aijmer et al. (2006).

LT: *Ji*   **juk**   *viską*   *daug*   *geriau*
      she.NOM   **juk**   everything.ACC   much   better

      supranta   *negu*   *sesuo.*
      understand.PRS.3   than   sister.NOM.SG[F]

      'She understands everything much better than the sister.'

(16) DE: *Laßt*   *mich*   *doch*   *zu*   *Gregor,*
          let.IMP.PL   I.ACC   doch   to   Gregor.DAT

          er *ist* ***ja*** *mein*   *unglücklicher*   *Sohn!*
          he is **ja** my.NOM.SG[M]   unlucky.NOM.SG[M]   son.NOM.SG[M]

      LT: *Leiskite*   *mane* *pas* *Gregorą,*   *jis*   ***juk*** *mano*
          let.2IMP.PL   I.ACC   to   Gregor.ACC[M]   he.NOM   **juk**   my

          nelaimingas   *sūnus!*
          unlucky.NOM.SG[M]   son.NOM.SG

          'Let me see Gregor! After all, he is my unlucky child.'[17]

I take this (at least occasional) "translatability" as the starting point and now turn to the German material.

German (as well as other West and North Germanic languages with the exception of English) exhibits a morphosyntactic class of elements which manifests properties similar to those I define as sentence particles in Lithuanian. In the German descriptive tradition, such elements are called *Abtönungspartikeln* ('nuance particles') or *Modalpartikeln*, and the labels *modal particles* or *discourse particles* are used in English-language publications. Modal particles constitute a clear-cut word class in German and some other West and North Germanic languages.[18] These are characterized by two main morphosyntactic properties:

– modal particles are bound uninflected forms,[19]

_____

[17] Lithuanian: http://www.tekstai.lt/buvo/versti/kafka/metamorf.htm. Accessed on 20.05.2019. German: https://www.bookrix.de/book.html?bookID=bibliothekar_1211810227.5542290211#0,558,41939. Accessed on 20.05.2019.

[18] English does not have a specific word class of this sort, being a remarkable exception among the Germanic languages.

[19] Some modal particles, e.g. *doch* and *ja* which are discussed below in more detail, have non-bound counterparts. Both *doch* and *ja* occur as independent utterances as response words ('quite the opposite' and 'yes', respectively). However, for descriptive purposes it is convenient to distinguish the "free" *doch* and *ja* from the homophonous modal particles, although they are clearly related historically.

– they occupy a strictly determined position in their host sentence (the so-called "middle field", "Mittelfeld"): a slot directly following the finite verb (or sometimes the direct object) in declarative sentences and the subject in interrogative sentences.

Additionally, modal particles are typically unstressed and integrated into the sentence intonation. Another non-trivial and typologically non-obvious characteristic of the Germanic modal particles is the "form – function" correlation. Indeed, modal particles that are found in the middle field tend to express the attitude of the speaker toward the proposition (Bross 2012, 185), whereas "connectives" or "conjunctions" whose functional domain is relating the proposition to the previous discourse unit usually occur sentence-initially (but there are exceptions). All these properties resemble those of the Lithuanian sentence particles, the main difference being the strictness of the position in a sentence (rigid in Germanic, more loose in Lithuanian).[20]

I will now focus on two German modal particles – *ja* and *doch* – whose functions, as I will show, are comparable to those of *juk* and *gi*. I now give a brief overview of how their functioning has been described in the literature.

From the very beginning, scholars noticed the functional affinity between *ja* and *doch*[21]. In fact, in many contexts they are interchangeable without a significant loss in sense. Consider an example in which the speaker starts a sentence with one particle, and then substitutes it with another after a pause in a self-repetition:

(17)   *… sie*   *hatten*   **doch**   *nur*   *eine,*   *die*
       they.NOM   have.PST.3PL   **doch**   only   one.ACC[F]   DEF.NOM.PL

       Römer   *hatten*   **ja** *nur*   *eine*   *Frau*   *gehabt.*
       Roman.NOM.PL   have.PST.3PL   **ja** only   one.ACC[F]   wife.ACC   have.PP

       'They used to have… The Romans used to have only one wife.' (Rath 1975, 237)

In earlier works, it was commonplace to claim that the core function of both *doch* and *ja* is to mark *mutual knowledge* or *shared knowledge* (Bross 2012, 197; Thurmair 1989, 104), or, formulated in a more formalized way, to relate a proposition to the *common ground*.[22] The notion "common ground" is defined in the following way:

"[the mutual beliefs] of the parties to a conversation are the beliefs they share, and that they recognize that they share: a proposition Φ is common belief of a group of

---

[20] See also Arndt (1960) for an interesting early comparison of the "modal particles" in German and Russian.

[21] For an overview of the affinities and differences between *ja* and *doch*, see Rinas (2006, 199–222).

[22] The term coined in Clark & Brennan (1991).

believers if and only if all in the group believe that Φ, all believe that all believe it, all believe that all believe that all believe it, etc." (Stalnaker 2002, 704)

As Bross (2012, 197) argues, this notion is problematic: a speaker cannot really know what the interlocutor knows or believes as we cannot see into each other's heads. In fact, both particles, especially *doch*, can be used when the hearer is clearly not aware of the proposition in advance:

(18) *Wir*   *sollten*   *den*   *Peter*   *einstellen.*   *Er*   *ist*
       we.NOM   must.PST.1PL   DEF.ACC.SG[M]   Peter.ACC   employ.PP   he.NOM   is

       doch   *ein*   *begabter*   *Junge.*
       doch   IND.NOM.SG[M]   talented.NOM.SG[M]   young.man.NOM.SG

       'We should have employed Peter. He is a talented young man.' (Rinas 2006, 205)

In (18), the speaker believes that the hearer knew but forgot Peter's virtues or even has an opposite opinion about him; the fact of Peter's being suitable for the job position is beyond doubt for the speaker her/himself. Therefore, less strong formulations are needed, e.g. as the following one: "[the] feature that *ja* and *doch* have in common is that in both cases the proposition expressed – say, P – is taken for granted by the speaker, and is assumed to be taken for granted by the hearer as well" (Gast 2008, 187). A good label that captures this is the notion of *uncontroversiality*, more precisely*, framing*[23] *the proposition as uncontroversial*. This term is used to some extent by Rinas (2006), and it is crucial in Grosz (2010).

Although *doch* and *ja* share the functional core of framing the proposition as uncontroversial and are often interchangeable, there are also differences between them. The most visible difference is the presence of an adversative meaning component in *doch* and its absence in *ja* (Thirmair 1989, 110–119; Grosz 2010). Grosz (2010, 163) also formulates this meaning component as *correction*. This can be seen in example (19):

(19)   a. *Ich*   *komme*   ***ja***   *schon.*
          I.NOM   come.PRS.1SG   **ja**   already

      b. *Ich*   *komme*   ***doch***   *schon.*
          I.NOM   come.PRS.1SG   **doch**   already

      'I am arriving.' (Rinas 2006, 200)

In (19a), the speaker expects the hearer to be aware of her/his arrival. In (19b), the speaker implies that his/her arrival is contrary to the expectations of the hearer. The

---

[23] I thank Michal Marmorstein (personal communication) for this formulation.

"correction" use may be seen as one of the manifestations of the contrastive component. In (20), the situation is described as follows: the hearer is amnesiac and forgot that s/he had been to Paris (e.g. says something like "I have never been to France"). The speaker corrects the hearer *reminding* him or her of the fact that this is not true:

(20) *Du*   *warst*   **doch** *schon*   *in* *Paris.*
      2SG.NOM   be.PST.2SG **doch** already in Paris.DAT

      '[Remember], you have been to Paris!' (Grosz 2010, 180)

The contrastive meaning component is most probably inherited from the source from which the modal particle *doch* evolved, namely the contrastive conjunction *doch* (a cognate of the English *though*) which is still in use in contemporary language, and occurs sentence-initially.[24]

On the other hand, *doch* is unable to occur in contexts expressing *general truth*, something which is presented as obvious and known to everybody. In such cases, only *ja* is possible:

(21) *Das*   *Thema*   *der*   *Vorlesung*   *sind*   *Katzen.*
      DEF.NOM.SG[N] topic.NOM.SG DEF.GEN.SG[F] lecture.GEN are   cat.NOM.PL

      Katzen   *gehören*   **ja/*doch** *zu*   *den*   *Säugetieren.*
      cat.NOM.PL belong.PRS.3PL ja   to   DEF.DAT.PL mammals.DAT.PL

      Alle   *Katzen*   *sind*   *Jäger.*
      all.NOM.PL   cat.NOM.PL   be.PRS.3PL   hunter.NOM.PL

      'The topic of our lecture is cats. [As you know], cats are mammals. All cats are hunters.' (Rinas 2006, 204)

Another important aspect in which *doch* and *ja* exhibit differences is their ability to occur in various types of speech acts. In this respect, *ja* turns out to be more restricted and only occurs in declaratives, see Table 2:

	declara-tive	polar interroga-tive	*wh*-inter-rogative	impera-tive	optative	exclama-tory	*wh*-ex-clamatory
doch	+		+	+	+		+
ja	+						

Table 2 (Thurmair 1989, 44). Occurrence of *ja* and *doch* in different speech acts

---

[24] For a brief note on the diachrony, see Conclusion.

On the basis of the analyses by Thurmair (1989), Karagjosova (2004), Rinas (2006), Gast (2008), and Grosz (2010), in Table 3 I provide a summary of the main contexts in which *doch* and *ja* occur ("functions")[25]. This is not exhaustive, but covers the most recurrent uses:

Functions/contexts	doch	ja	
Uncontroversial information (declaratives)	+	+	1
Uncontroversial information + correction (declaratives)	+		2
Uncontroversial information + reminding (declaratives)	+		3
Uncontroversial information + general truth		+	4
Wh-Exclamations	+		5
Wh-Questions	(+)[26]		6
Tag questions (positive response expected)	+		7
Imperatives	+		8

Table 3. The main occurrence contexts of *doch* and ja in German[26]

3.2 *Juk* and *gi* functioning in Lithuanian: A revision

Having in mind the portrait of the German particles *doch* and *ja* sketched in the previous section, I will now turn back to the Lithuanian *juk* and *gi*. Their morphosyntactic characteristics have already been addressed in Section 2.2, and now I will discuss their contribution to meaning.

It is clear that many German sentences presented in the previous section can be translated into Lithuanian using either *juk* or *gi*, or both. Let us now consider some of their core uses.

As *doch* and *ja*, *juk* and *gi* seem to be interchangeable in many contexts. If one translates the German example (17) into Lithuanian, it is clear that both *juk* and *gi* are perfectly acceptable:

(22) a. ***Juk***[27] *romėnai*   *turėjo*   *tik*   *vieną*   *žmoną.*
         juk   Roman.NOM.PL have.PST.3   only   one.ACC.SG[F]   wife.ACC.SG

      b. *Romėnai*   ***gi***   *turėjo*   *tik*   *vieną*   *žmoną.*
         Roman.NOM.PL **gi**   have.PST.3   only   one.ACC.SG[F]   wife.ACC.SG

      'Romans used to have only one wife.'

---

[25] In this paper, I use the terms "context" and "function" as technical terms which are synonyms. To get the idea of what kinds of entities such contexts/functions are, I redirect the reader to the works written in a similar spirit: Dahl (1985), Haspelmath (1997), van der Auwera & Sahoo (2015).

[26] The use of *doch* is reported to be somewhat marginal in *wh*-interrogatives. For discussion, see Section 3.3.

[27] The initial position of *juk* is default, but second position is also acceptable.

The main difference between (22a) and (22b) reported by Lithuanian speakers is that (22a) sounds more formal, whereas (22b) is more colloquial. However, if one includes both sentences in a broader context, it turns out that *juk* and *gi* do not sound equally well. Imagine the following situation. The speaker is arguing that the social position of woman in the Roman society was relatively high, and this was related to the fact that Romans used to have only one wife. In this context, both (22a) and (22b) are equally acceptable. By contrast, if the interlocutor asks something like "How many wives did Caesar have?", the variant with *gi* (22b) sounds better than (22a). With *gi* the speaker not only indicates that there is the uncontroversial fact that Romans used to have only one wife, but also that, in her/his view, there are inconsistencies in the hearer's beliefs. This "contrastive" component of the meaning of *gi* is clearly parallel to that of the German *doch* (functions 2-3 from Table), whereas its absence is comparable to that of the German *ja*, although it seems that in Lithuanian this contrast is less strict: *juk* is still possible in contrastive contexts. The "correction" function of *gi* is even better observable in (23), which presents a fragment of a dialogue between a child and an adult:

(23)   – *Kokia   didelė   gorila!*
          what.NOM.SG[F]   big.NOM.SG[F]   gorilla.NOM.SG

        – *Čia   gi   meška!*[28]
          here gi   bear.NOM.SG

        '– What a huge gorilla! – But it's a bear!'

Speakers report that in (23) both *gi* and *juk* are possible, *gi* being, however, more natural, especially in colloquial speech.

The "reminding" function is illustrated in (24). The hearer seems to be surprised by the fact that it is early evening, and alcohol is not being sold in the supermarket. The speaker reminds her/him:

(24)   *Šiandien gi sekmadienis!*
        today   gi Sunday.NOM.SG

        '[Of course], today is Sunday'

---

[28] The present example is also remarkable in the sense that it exhibits an interesting interaction between the presence/absence of a copula and the presence/absence of a sentence particle. Some speakers report that either *Čia yra* [COP] *gorilla* or **Juk** *čia gorila* / *Čia* **gi** *gorilla* sound natural, but not ?**Juk** *čia yra gorila* / *Čia* **gi** *yra gorila*. This shows that the sentence particles in Lithuanian may have some copular properties. This has been noticed by Wiemer (2007, 178) for the evidential marker *esą,* which has a lot in common with *juk*, *gi* and *ar* morphosyntactically but is more autonomous (e.g. may be followed by a pause). Typologically, an interesting parallel is Sanzhi Dargwa (Forker, forthcoming) where the copula is not compatible with =*q'al* (the Sanzhi Dargwa functional parallel to *juk* and *gi*).

In (24), the speaker appeals to the supposed background knowledge of the hearer: alcohol is not sold in Lithuanian supermarkets after 3 p.m. on Sundays. In this example, *gi* is reported to sound better than *juk,* although *juk* is not completely excluded. *Juk* is more natural if the hearer is not expected to disagree and is supposed to be aware of both facts: (1) alcohol is not sold in Lithuanian supermarkets after 3 p.m. on Sundays, (2) today is Sunday. Imagine a situation in which two friends have a plan to have a beer on Sunday after 3 p.m. Theoretically, they have a choice: to buy beers in a supermarket and drink them at home or to go to a bar; in fact, only one option is realistic. The speaker A supposes the speaker B to be aware of both (1) and (2) and says:

(25) *Eime* *į* *barą.* *Šiandien* **juk** *sekmadienis.*
go.IMP.1PL in bar.ACC.SG today **juk** Sunday.NOM.SG
Let's go to a bar. It's Sunday'

In (25), both *juk* and *gi* are reported to sound equally natural (as in other cases, *gi* being more colloquial). The contrastive component is lacking here.

More visible and strict differences between *juk* and *gi* concern their ability to occur in non-declarative sentences. Only *gi* occurs in *wh*-questions and *wh*-exclamatives:

(26) *O* *kas* **gi** *jiems* *suteikė* *monopolistų*
but who.NOM **gi** they.DAT[M] give.PST3 monopolist.GEN.PL
teisę?
right.ACC.SG
'Then, who gave them the monopoly right?' (LKT)

(27) *Kaip**gi**[29]* *gražiai* *ji* *šoka!*
how.GI beautiful.ADV she.NOM dance.PRS.3
'Isn't it beautiful how she is dancing?'

In both cases, the use of *juk* would be impossible. However, *juk* can appear in questions if its scope is broader than the question itself, like (28) where the rhetorical question is "indented" within an intended declarative sentence:

(28) *O* **juk** *kas* *jiems* *suteikė* *monopolistų* teisę*? – [Jis!]*
but **juk** who.NOM they[M].DAT give.PST3 monopolist.GEN.PL right? he.NOM
'Who gave them the monopoly rights? [He did].'

[29] The joint or separate spelling of *gi* is conventional, explained by the history of normalization of the Lithuanian spelling system.

In (28), *juk* takes scope over the intended proposition <It was he who gave them the monopoly rights> and not over the question only, which is rhetorical, and frames the whole intended proposition expressed by two sentences as uncontroversial.

It is not easy to explain how the marking of uncontroversial information is semantically related to the use of *gi* in questions. However, there are reasons to claim that *gi* only occurs in questions whose content is framed as expected by the speaker, consider example (29):

(29) *Kas   gi   laimėjo   rinkimus?*
      who.NOM **gi** win.PST.3   election.ACC.PL

      'So, who won the elections?'

This sentence sounds natural if pronounced in the situation in which everybody is waiting for the results of the elections. By contrast, it can hardly be pronounced when the talk is about the weather, and, suddenly, someone has the idea to ask about the elections' results. In the latter case, the sentence would sound more natural without *gi* (and would possibly start with *beje* 'by the way').

Both *gi* and *juk* occur in the questions equivalent with tag questions in English when the expected response is positive:

(30) *Šiandien* **juk/gi**   *sekmadienis?*
      today   **juk/gi**   Sunday.NOM.SG

      'Todays is Sunday, isn't it?'

The combination of the polar question particle *ar* and *gi* (spelled as a single word *argi*) mark polar questions in cases when the speaker has doubts about the truth of the proposition under the scope of the question marker:

 (31) *Ar***gi**   *šiandien*   *sekmadienis?*
      Q.**gi**   today   Sunday.NOM.SG

      'Is it really Sunday today?'

Finally, *gi* but not *juk* is used with imperatives in colloquial language. Such imperatives mark the impatience of the speaker who expects the hearer to be aware of the speaker's wish but not to be acting for some reason as in (32):

(32) *Įpilk*   *gi*   *ir*   *man*   *arbatos.*
      pour.IMP.2SG   **gi** and   I.DAT   tea.GEN.SG

      '[Come on], pour me some tea too!'

In Table 4, I summarize the prominent contexts/functions in which the Lithuanian sentence particles *juk* and *gi* occur. As in the case of German (Table 3), these contexts are not exhaustive, though they do reflect the general picture and enable cross-linguistic comparison. (–) means that the use is allowed but dispreferred / not very natural.

Functions/contexts	gi	juk	
Uncontroversial information (declaratives)	+	+	1
Uncontroversial information + correction (declaratives)	+	(–)	2
Uncontroversial information + reminding (declaratives)	+	(–)	3
Wh-Exclamations	+		4
Wh-Questions	+		5
Tag questions (positive response expected)	+	+	6
Tag questions (negative response expected)	+		7
Imperatives	+		8

Table 4. The main occurrence contexts of *gi* and *juk* in Lithuanian

3.3 German *vs* Lithuanian

In the previous section, I analyzed the sentence particles *juk* and *gi* of Lithuanian using some insights gained in the studies of German modal particles *doch* and *ja*. I now provide a brief comparison of all four particles.

Concerning the positions in sentence of *juk* and *gi*, on the one hand, and *doch* and *ja*, on the other hand, these are defined on the basis of language-specific rules in both German and Lithuanian. These rules are different, and the German particles are characterized by stricter positioning (in the "middle field").

The functions of all four particles are summarized below in Table 5.

Functions/contexts	gi	juk	doch	ja	
Uncontroversial information (declaratives)	+	+	+	+	1
Uncontroversial information + correction (declaratives)	+	(–)	+		2
Uncontroversial information + reminding (declaratives)	+	(–)	+		3
Uncontroversial information + general truth				+	5
Wh-Exclamations	+		+		5
Wh-Questions	+		(+)		6
Tag questions (positive response expected)	+	+	+		7
Tag questions (negative response expected)	+				8
Imperatives	+		+		9

Table 5. *Doch*, *ja*, *juk* & *gi*: comparison

Indeed, all four particles of the two languages are characterized by the common meaning component which has been previously defined as framing the proposition as uncontroversial. German and Lithuanian exhibit two particles each, and in both languages there are contexts in which the particles are interchangeable. Moreover, in both German and Lithuanian, there are particles with broader and narrower functional scales (*doch & gi* vs *ja & juk*, respectively). In Lithuanian, as well as in German, the particle with a broader functional scale is characterized by the additional *contrastive* meaning component (*doch* and *gi*). For both particles, the presence of this component correlates with the ability to occur in non-declarative sentences.

Regarding the latter aspect, *doch* and *gi* also exhibit differences: *gi* occurs in non-declarative contexts more easily than *doch*. *Doch* does not occur in tag questions of the type 8. In such cases, a reformulation is needed (33a), or another particle, *denn* (33b), is used:

(33) a. *Ist   heute   wirklich   Sonntag?*
   is today really Sunday.NOM.SG

  b. *Ist   heute   denn   Sonntag?*
   is today really Sunday

   'Is it really Sunday today?'

Moreover, unlike *gi*, the German *doch* occurs in *wh*-questions only marginally. The most typical occurrence of *doch* in *wh*-questions in German are the situations in which the speaker forgot something and wants the hearer to help recover the knowledge. It often co-occurs with another particle *gleich*:

(34) *Wer   war   das **doch** gleich?*
  who.NOM was it **doch** gleich

  'Who was it?'

In (34), the speaker expects her/himself to know who the person was who has just passed by, but suddenly s/he realizes s/he has forgotten the name.

In the cases like the Lithuanian (29), when the speaker asks a question with an expected topic, German normally uses *denn*:

(35) *Wer   hat   **denn** die   Wahl   gewonnen?*
  who.NOM have.PRS.3 **denn** the.ACC[F] elections.ACC won

  'So, who won the elections?'

To summarize, all the four particles of both languages exhibit a common functional core. Roughly speaking, the functional scale of *doch* is very much like the one of *gi*, whereas *ja* exhibits similarities with *juk*. In both cases, the coincidence is not complete, but there are more similarities than differences. It is also to be noted that the distribution of *juk* and *gi* in Lithuanian is freer than that of *doch* and *ja*: the Lithuanian particles are interchangeable in more contexts.

4 Methodological, areal, and diachronic dimensions

4.1 Methodology: language-particular *vs* cross-linguistic

The methodological stance invoked in this study – using insights gained in descriptions of one language in order to understand structures and functions of another – refers to the current debate in linguistic typology known as the 'comparative concepts' *vs* 'descriptive categories' problem. The debate was triggered by Martin Haspelmath at the end of the previous decade, and is still ongoing. In Haspelmath's (2010) view, particular languages should be described *in their own terms* (Boas 1911) in a structuralist/descriptivist spirit. *Descriptive categories* are always *language-particular*, and languages are best compared through *comparative concepts* consciously designed by linguists. The latter are clearly formulated, uncontroversial *definitions* of compared phenomena made up by linguists, and not *pre-established categories* or *cross-linguistic categories*, which, supposedly, do not exist as *natural kinds* (Haspelmath 2007, 2018). For now, one can say that Haspelmath's main distinction between comparative concepts and descriptive categories has been generally accepted by the typology community.

However, this radical stance has also been criticized. Many points of criticism are collected in a special issue of *Linguistic Typology* (20(2), 2016). The main argument of most of Haspelmath's critics is that comparison between languages and the knowledge of cross-linguistic patterns is doubtlessly useful in discovering language-particular structures. As Lander and Arkadiev (2016, 412) put it, "being informed about other languages and typological variation [...] in general enables one to [...] see systematic patterns in the data which otherwise might appear just chaotic."

In this paper, I have followed the spirit of the study by van der Auwera & Sahoo (2015), which is one of the most influential papers criticizing Haspelmath's approach. Arguing against his extreme *categorical particularism* (though accepting it in a softer version), the authors demonstrate that the distributional and functional properties of English *such* can be better understood if compared to similar elements of other languages. It turns out that the functional core of these elements becomes more visible in cross-linguistic comparison: all the languages mentioned in the paper exhibit elements with a non-trivial set of common properties, being more similar to each other than to anything else in their

own languages. This enables the authors to propose a functional category ("similative") which describes to a large extent the behavior of different language-particular elements.

In this study, I have demonstrated how an analysis originally designed for one language can be fruitfully applied to another language, very much as van der Auwera & Sahoo (2015) did. In doing so, however, it was important to keep in mind that descriptive categories of one language cannot be mechanically transferred onto another: commensurable language-particular categories always exhibit similarities as well as differences. These can be formalized in describing their functioning in the form of sets of elementary uses/functions, as presented in Table 5. The data presented in Table 5 is a sort of comparative concept designed on the basis of only two languages – German and Lithuanian. On the other hand, this comparative concept gives one strong descriptive instruments for research on other languages.

4.2 Areal-typological and diachronic dimensions

German and Lithuanian are not located on different continents; both are Indo-European, both belong to the European macroarea (Haspelmath 2001), but also to the smaller area adjacent to the Baltic Sea (Koptjevskaja-Tamm and Wälchli 2001). They are historically interconnected, and their structures are definitely commensurable, although the degree of direct influence of one upon another is disputable. Although large-scale cross-linguistic studies of similar categories have not yet been published,[30] there are multiple language-particular descriptions in which some elements are described in very similar terms. Many such languages belong to the same Eastern, Central, and Northern European area. Obvious parallels to the particles discussed in the present paper are the Russian *že* and *ved'* (see Valova (2016, 85–102) for an overview) and their parallels in Ukrainian (*-ž*, *žež*, *adže*) and Belarusian (*ž*[*a*]), the North Russian dialectal *dak* (Post 2006), the particles *ju/jo* of the Scandinavian languages (Andvik 1994; Aijmer 1996), the *-han/-hän* "clitic" of Finnish (Palomäki 2009). There are some contrastive studies between pairs of languages: German *vs* Russian (Orlova 2012), German *vs* Dutch (Foolen 2006), German *vs* Czech (Rinas 2006). In the parallel Lithuanian-Latvian corpus, *gi* and *juk* regularly correspond to the Latvian *taču*.[31] Beyond this part of Europe, similar elements are reported for the Northeast Caucasian (Nakh-Daghestanian) languages: Standard Lezgian (Haspelmath 1993, 242) and Sanzhi Dargwa (Forker, forthcoming)[32].

---

[30] Panov (forthcoming) focuses on the function analyzed here for Lithuanian in a broader typological and areal perspective.

[31] http://tekstynas.vdu.lt/page.xhtml?id=parallelLILA. The parallel Latvian-Lithuanian subcorbus of LKT (8 782 050 words). Accessed on 24.05.2019.

[32] In Sanzhi Dargwa, the element =*q'al* is clearly a verbal affix as it is obligatorily adjacent to a verb (Forker, forthcoming): that is, =*q'al* is morphosyntactically distinct from the sentence particles of Lithuanian and the modal particles of German.

Another aspect relevant for understanding the typology and areality of the elements framing proposition as uncontroversial is their diachrony. Regarding the particles considered in the present paper, the diachronic path is clearer for German than for Lithuanian. The German modal particle *ja* was already present in Old High German (spelled as *ia*) in a function close to the contemporary one (Wauchope 1991, 93–131), but it used to occur in positions other than the "middle field". It had originated form the homophonous word meaning 'yes'[33] which is reconstructed for Proto-Germanic (Kluge 1989[1883], 338) and survives in contemporary German. The particle *thoh* (>*doch*) also functioned in Old High German in the way resembling the contemporary one, and it derives from the homophonous adverb/conjunction which has a contrastive meaning ('though, however, but')[34].

Both *juk* and *gi* are attested already in the earliest Lithuanian written texts, but their etymology remains somewhat obscure. *Juk* also occurs as *jukaigi*, *jukaig*, *jukaigei*, *jukag*, *jukig* in Old Lithuanian (Smoczyński 2017, 527), and its function seems identical to the contemporary one; consider an example from Daukša's Catechism (1599):

(36)   *Bęt' ką*   *dauġ* *kałbét'?* *Iuk'* *ir*   *patis*
      but what.ACC much speak.INF juk and self.NOM.SG(M)

      Arciheretikas   *Martinas*   *Luthęris*   *ſawamé*
      grand.heretic.NOM.SG Martin.NOM.SG Luther.NOM.SG REFL.POSS.LOC.SG[M]

      Catechiſmé   *rágina*   *iſkáłos*   *Miſtrús*
      catechism.LOC.SG admonish.PRS.3 school.GEN.SG teacher.ACC.PL

      idąnt' pratintų   *waikús*   *ſawús.*
      that   accustom.SBJ.3 children.ACC.PL REFL.POSS.ACC.PL[M]

     'But why speak more? After all, the grand heretic, Martin Luther, in his […] catechism admonishes school masters, that they accustom their children [to make the sign of the cross].' (Daukša 2000, 959, lines 4-6, quoted from Ostrowski 2015, 202)

The etymology of *juk* is indicated as obscure by Smoczyński (2017, 453). Ostrowski (2015) suggests deriving it from the combination of *juo* 'especially, notably, all the more' and *kai* 'when'. The path of development of the particle *gi* is even more obscure. In Old Lithuanian, the particle clearly had a phrasal, not sentential scope, probably marking some kind of focus as in the *Postilė* by Bretkūnas:[35]

---

[33] But see Petrova (2016) for the overview of an alternative hypothesis which derives the contemporary modal particle *ja* form the adverb *je* 'always'.

[34] A similar pathway ('simple contrast' > 'the marking of uncontroversial information') is to be suspected at least in the cases of East Slavic (*že*) and Latvian (*taču*). More cross-linguistic research is needed here, see also Panov (forthcoming).

[35] For an overview of the previous studies and multiple examples, see Ambrazas (2006, 80–82).

(37) *Potam*   *ischgulda*   *iemus*   *Penktan-**gi***
      then   explain.PRS.3   they.DAT.PL[M]   fifth.ACC.SG-**gi**

      prisakima   *sawa*   *dangaus*   *Tiewa.*
      commandment.ACC.SG   RFL.POSS.ACC.SG   heaven.GEN.SG   father.GEN.SG

      'Then he lectured them on the fifth commandment of his heavenly Father'
      (Bretkūnas 2005, quoted from Nau and Ostrowski 2010, 26)

Etymologically, *gi* derived from either the **g-* or the **gh-* particle of Proto-Indo-European, possibly from both particles which merged phonetically. The original meanings of **g-* or **gh-* remain relatively obscure, and Common Slavic **že*, and both Greek *de* and *ge* are possible cognates (Dunkel 2014, 279–282). However, it is unclear how exactly *gi* came to function the way it does in contemporary Lithuanian, and more research is needed. It is clear that the particles of Lithuanian and German grammaticalized from different sources in different periods. More areal-typological as well as historical data is required to understand how different developments led to similar results in the two languages discussed in this paper.

5 Conclusion

In this study, I investigated the elements of contemporary Lithuanian known as "particles" in traditional descriptions. I demonstrated that elements labelled in this way actually exhibit very diverse morphological and syntactic properties. This particularly concerns the issues of boundness and scope relations. Lithuanian "particles" can take scope over a single word, a phrase, or a sentence/utterance. I suggested labelling the latter type *sentence particles*. I then focused on a case study, namely two sentence particles – *juk* and *gi* – which, as I argue, belong to this category, and whose functions had not been properly investigated. Invoking some influential descriptions of German particles *ja* and *doch*, I argued that *juk* and *gi* can be analyzed in the same terms, namely, as markers whose core function is *framing the proposition under scope as uncontroversial*. More than that, *juk* and *ja*, on the one hand, and *gi* and *doch*, on the other, exhibit subtler functional parallels in that the latter two occur in contrastive contexts, *wh*-questions, *wh*-exclamatives, and commands, whereas the former are restricted to statements and do not imply a contrast.

Acknowledgements

Abbreviations & glosses

1 = first person; 2 = second person, 3 = third person; ACC = accusative; ADV = adverb; AOR = aorist; GEN = genitive; DAT = dative; DEF = definite article; F = feminine; FT = future; IMP = imperative; IND = indefinite article; INF = infinitive; LKT = Lietuvių kalbos tekstynas [The corpus of Lithuanian]; LOC = locative; M = masculine; NEG = negation; NOM = nominative; PL = plural; PP = past participle; PRS = present; PST = past; Q = question particle; SBJ = subjunctive; SG = singular.

Data source

LKT: Dabartinės lietuvių kalbos tekstynas. Corpus of the Contemporary Lithuanian Language. 2011. Available at: http://tekstynas.vdu.lt/tekstynas.

References

Aijmer, Karin, Ad Foolen & Anne-Marie Simon-Vandenbergen. 2006. Pragmatic markers in translation: a methodological proposal. In *Approaches to Discourse Particles*. Kerstin Fischer, ed. Amsterdam: Elsevier. 101–114. https://doi.org/10.1075/bct.44.09han.

Aijmer, Karin & Anne-Marie Simon-Vandenbergen, eds. 2006. *Pragmatic Markers in Contrast*. Amsterdam: Elsevier.

Ambrazas, Vytautas, ed. 1997. *Lithuanian Grammar*. Vilnius: Baltos lankos.

Ambrazas, Vytautas. 2006. *Lietuvių kalbos istorinė sintaksė.* [The Historical Syntax of Lithuanian]. Vilnius: Lietuvių kalbos institutas.

Andvik, Erik E. 1992. *A Pragmatic Analysis of Norwegian Modal Particles*. Arlington: Summer Institute of Linguistics & University of Texas at Arlington.

Arndt, Walter. 1960. "Modal particles" in Russian and German. *Word* 16(3), 323–336.

Bloomfield, Leonard. 1933. *Language*. New York: H. Holt and Company.

Boas, Franz. 1911. Introduction. *Handbook of American Indian Languages, vol. 1 (Bureau of American Ethnology, bulletin 40)*. Washington, DC: Government Printing Office.

[Bretkūnas, Jonas]. 2005. *Jono Bretkūno Postilė: Studija, faksimilė ir kompaktinė plokštelė*. Lietuvių kalbos institutas. Edited by Ona Aliknavičienė. Vilnius: Lietuvių kalbos instituto leidykla.

Bross, Fabian. 2012. German modal particles and the common ground. *Helikon* 2, 182–209.

[Daukša, Mikalojus]. 2000. *Mikalojaus Daukšos 1599 metų Postilė ir jos šaltiniai*. Edited by Jonas Palionis. Vilnius: Baltos lankos.

Clark, Herbert H. & Susan E. Brennan. 1991. Grounding in Communication. In *Perspectives on socially shared cognition*. Lauren B. Resnick, John M. Levine

& Stephanie D. Teasley, eds. Washington: American Psychological Association. 127–149. https://doi.org/10.1037/10096-006.

Degand, Liesbeth, Bert Cornillie & Paola Pietrandrea, eds. 2013. *Discourse Markers and Modal Particles: Categorization and Description*. Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1177/1461445615588172b.

Dunkel, George E. 2014. *Lexikon der indogermanischen Partikeln und Pronominal-stämme*. Band 2: Lexikon. Heidelberg: Universitätsverlag Winter.

Fedriani, Chiara & Andrea Sansò, eds. 2017. *Pragmatic markers, discourse markers, and modal particles*. Amsterdam & Philadelphia: John Benjamins. https://doi.org/10.1075/slcs.186.

Foolen, Ad. 2006. Polysemy Patterns in Contrast: The Case of Dutch *toch* and German *doch*. In *Pragmatic Markers in Contrast*. Karin Aijmer & Anne-Marie Simon-Vandenbergen, eds. Amsterdam: Elsevier. 59–72. https://doi.org/10.1075/lic.7.2.12and.

Forker, Diana, *forthcoming*. More than just a modal particle: The enclitic *=q'al* in Sanzhi Dargwa. To be published in *Functions of Language*.

Gast, Volker. 2008. Modal particles and context updating – the functions of German *ja*, *doch*, *wohl* and *etwa*. In *Modalverben und Grammatikalisierung*. Heinz Vater & Ole Letnes, eds. Trier: Wissenschaftlicher Verlag. 153–177.

Grosz, Patrick. 2010. German *doch*: An element that triggers a contrast presupposition. *Proceedings of the Chicago Linguistic Society* 46, 163–177.

Haspelmath, Martin. 1993. *A Grammar of Lezgian*. Berlin, New York: Mouton de Gruyter.

Haspelmath, Martin. 2001. The European linguistic area: Standard Average European. In *Language Typology and Language Universals. Sprachtypologie und sprachliche Universalien. Sprachtypologie und sprachliche Universalien: La typologie des langues et les universaux linguistiques: An International Handbook: Ein internationales Handbuch: Manuel international*. Martin Haspelmath, Ekkerhard König, Wolfgang Oesterreicher & Wolfgang Raible, eds. Berlin & New York: Walter de Gruyter. 1492–1510.

Haspelmath, Martin. 2007. Preestablished categories don't exist: Consequences for language description and typology. *Linguistic Typology* 11(1), 119–132. https://doi.org/10.1515/lingty.2007.011.

Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3), 663–687. https://doi.org/10.1353/lan.2010.0021.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45(1), 31–80. https://doi.org/10.1515/flin.2011.002.

Haspelmath, Martin. 2013. Argument indexing: A conceptual framework for the syntactic status of bound person forms. In *Languages across boundaries*. Dik Bakker & Martin Haspelmath, eds. Berlin: Mouton de Gruyter. 209–238. https://doi.org/10.1515/9783110331127.197.

Haspelmath, Martin. 2018. How comparative concepts and descriptive linguistic categories are different. In *Aspects of Linguistic Variation*. Daniël Olmen, Tanja Mortelmans & Frank Brisard, eds. Berlin, Boston: Mouton de Gruyter. 83–113. https://doi.org/10.1515/9783110607963-004.

Karagjosova, Elena. 2004. *The Meaning and Function of German Modal Particles*. PhD thesis. Berlin: Freie Universität.

Kluge, Friedrich. 1989. *Etymologisches Wörterbuch der deutschen Sprache*. 22. Auflage. Berlin & New York: Walter de Gruyter.

Koptjevskaja-Tamm, Maria & Bernhard Wälchli. 2001. The Circum-Baltic languages: An areal typological approach. In *Circum-Baltic Languages. Volume 2: Grammar and Typology*. Östen Dahl & Maria Koptjevskaja-Tamm, eds. Amsterdam & Philadelphia: John Benjamins. 615–750. https://doi.org/10.1075/slcs.55.15kop.

Kornfilt, Jaklin. 1997. *Turkish*. London & New York: Routledge.

Lander, Yury & Peter Arkadiev. 2016. On the right of being a comparative concept. *Linguistic Typology* 20(2), 403–416.

Munaro, Nicola & Cecilia Poletto. 2004. Sentential Particles and Clausal Typing in the Veneto Dialects. In *Proceedings of the Dislocated Elements Workshop : ZAS Berlin, November 2003*. Benjamin Shaer, ed. Berlin: Zentrum für Allgemeine Sprachwissenschaft, Sprachtypologie und Universalienforschung. 375–397.

Nau, Nicole & Peter Arkadiev. 2015. Towards a standard of glossing Baltic languages: The Salos Glossing Rules. *Baltic Linguistics* 6, 195–241. https://doi.org/10.32798/bl.398.

Nau, Nicole & Norbert Ostrowski. 2010. Background and perspectives for the study of particles and connectives in Baltic languages. In *Particles and connectives in Baltic* (Acta Salensia 2). Nau, Nicole & Norbert Ostrowski, eds. Vilnius: Vilniaus Universitetas. 1–38.

Ostrowski, Norbert. 2015. The Origin of the Lithuanian Particle »jùk«. In *Proceedings of the 4th Greifswald University Conference on Baltic Languages*. Artūras Judžentis & Stefan Kessler, eds. Berlin: Logos Verlag. 201–215.

Palomäki, Jennimaria Kristiina. 2009. *The Pragmatics and Syntax of the Finnish -han Particle Clitic*. PhD thesis. Athens, GA: The University of Georgia.

Panov, Vladimir A. *Forthcoming*. The marking of uncontroversial information in Europe: Presenting the enimitive. To appear in *Acta Linguistica Hafniensia*.

Petrova, Svetlana. 2016. On the status and the interpretation of the left-peripheral sentence particles *inu* and *ia* in Old High German. In *Discourse Particles: Formal Approaches to their Syntax and Semantics*. Josef Bayer & Volker Struckmeier, eds. Berlin & Boston: de Gruyter. https://doi.org/10.1515/9783110497151-012.

Post, Margje. 2005. *The Northern Russian pragmatic particle dak in the dialect of Varzuga (Kola Peninsula)*. PhD thesis. Tromsø: Tromsø Universitet.

Rath, Rainer. 1975. *Doch*. Eine Studie zur Syntax und zur kommunikativen Funktion einer Partikel. *Deutsche Sprache* 3, 222–242.

Rinas, Karsten. 2006. *Die Abtönungspartikeln doch und ja: Semantik, Idiomatisierung, Kombinationen, tschechische Äquivalente*. Frankfurt am Main: Peter Lang. https://doi.org/10.1515/zrs.2010.025.

Smoczyński, Wojciech. 2017. *Lithuanian etymological dictionary*. Berlin: Peter Lang.

Stalnaker, Robert. 2002. Common Ground. *Linguistics and Philosophy* 25, 701–721.

Thurmair, Maria. 1989. *Modalpartikeln und ihre Kombinationen*. Tübingen: Max Niemeyer Verlag. https://doi.org/10.1515/9783111354569.

Valova, Evdokia A. 2016. *Pozicionnye svojstva enklitičeskich častic: korpusnoe i eksperimental'noe issledovanie na primere časticy že.* [Positional properties of enclitic particles: a corpus and expreimantal study based on particle že]. A PhD *thesis*. Moskva: National Research University "Higher School of Economics".

van der Auwera, Johan & Kalyanamalini Sahoo. 2015. On comparative concepts and descriptive categories, such as they are. *Acta Linguistica Hafniensia* 47(2), 136–173. https://doi.org/10.1080/03740463.2015.1115636.

Wauchope, Marie Michele. 1991. *The Grammar of the Old High German Particles* thoh*, ia*, and thanne. New York, Bern, Frankfurt/M., Paris: Peter Lang.

Weidt, Harald, ed. 1989. *Sprechen mit Partikeln*. Berlin & New York: Mouton de Gruyter.

Wiemer, Björn. Lexical markers of evidentiality in Lithuanian. *Rivista di Linguistica* 19(1), 173–208.

Zaliznjak, Andrej A. 2008. *Drevnerusskie ènklitiki.* [Enclitics in Old Russian]. Moskva: Jazyki slavjanskich kul'tur.

Zimmerling, Anton V. 2013. *Sistemy porjadka slov v slavjanskich jazykach v tipologičeskom aspekte.* [Slavic word order systems in a typological perspective]. Moskva: Jazyki slavjanskoj kul'tury.