

Complex Sentences and their Punctuation in English Texts Composed by Latvian Students

Zigrīda Vincēla

University of Latvia Faculty of Humanities, Visvalža St. 4A, Riga,
LV-1050, Latvia, zigrida.vincela@lu.lv

Summary. Comma use in Latvian and English complex sentences differs significantly and this can cause major problems to Latvian users of English. In particular, the dependent clause that follows the independent clause is normally separated by a comma in Latvian complex sentences (Blinkēna, 2009), whereas comma use/non-use in English complex sentences (Downing, Locke, 2006) depends on the semantic relation between the independent and the dependent clause. These comma use differences, as previous research reveals (Farneste, 2006b), can cause challenges in English complex sentence punctuation to students whose native language is Latvian.

The goal of this study is to research Latvian students' punctuation of English complex sentences extracted from a corpus of essays. The analysis covers the cases of nominal and relative dependent clauses that are tightly related to the independent clause they follow. Normally, as Downing and Locke (2006) note, such dependent clauses are not separated by a comma in English texts. The analysis shows that a strong semantic relation between the clauses has been relevantly recognised by students in numerous examples, which they have relevantly punctuated. However, the results also reveal unnecessary separations of relative and nominal clauses from the independent clause by a comma. These findings call for the focusing of students' attention on the role of the semantic relation of clauses in English complex sentence punctuation by contrasting these sentences with the relevant complex sentences found in Latvian texts. In addition, further analysis of various sentence types and their punctuation in student-composed texts of different communicative purposes is required.

Keywords: *complex sentence, clause, punctuation, semantic dependency.*

Introduction

Comma use in Latvian and English complex sentences differs significantly and this can cause major problems to Latvian users of English. The dependent clause that follows the independent clause is normally separated by a comma in Latvian complex sentences (Blinkena, 2009), whereas comma use/non-use in English complex sentences (Downing and Locke 2006) is related to the semantic interdependency between the independent and the dependent clause. These and other differences in comma use have been previously discussed by Farneste (2006a, 2006b). She concludes that differences can cause English sentence punctuation challenges to students whose native language is Latvian. Farneste (2006a, 37) emphasises that first-year university students in Latvia, when asked if they discussed punctuation in English texts in English classes at their schools, predominantly point out that it was rarely done. The author also notes that there are very few tasks in English course books focusing on the communicative function of punctuation marks. Therefore, students often tend to rely on the punctuation rules that have been discussed in Latvian language classes. Meanwhile, the research on comma overuse in complex sentences in English texts written by Latvian students has not been taken up yet.

The goal of this study is to explore examples of complex sentences written by Latvian students in order to uncover how far their punctuation corresponds to the semantic relation between the independent and dependent clauses.

Seeking to achieve this goal, the following enabling objectives are set: (1) a comparative analysis of the theories on sentence classification according to clause relations and clause types; (2) extraction of examples of complex sentences from a corpus of essays composed by Latvian students (with the help of *AntConc*) that have been part-of-speech tagged with CLAWS7; (3) grouping of the extracted examples into sentences that contain nominal and relative dependent clauses; (4) analysis of the semantic relation of the dependent and independent clauses and their punctuation as proposed by the students.

Theoretical background

Sentences, 'the highest grammatical units' (Downing, Locke, 2006, 272), that contain more than one clause or verb phrase (Biber et al., 1999, 100, 120) as a central element are classified by linguists in terms of the relationship of the elements within the sentence. Researchers of the Latvian language (e.g. Auziņa et al., 2013, 831; Beitiņa, 2009, 173; Ceplītis et al., 1989, 110–111) label all such sentences as composite sentences (in the Latvian language: *salikti teikumi*) and use this as an umbrella term for all sentences that go beyond one clause. This term is not normally found in the grammar

of the English language. However, it has been used by some linguists, for example, Lyons (1995, 157) in order to distinguish ‘non-simple’ sentences, which contain more than one clause, from simple sentences. It is also used by Davidson (2005, 30) in his punctuation manual ‘*How to Punctuate*’.

Researchers of the Latvian language and other linguists (e.g. Auziņa et al., 2013, 832; Ceplītis et al., 1989, 110–111; Downing, Locke, 2006, 272; Yule, 2006, 266) distinguish compound sentences (in the Latvian language: *salikti sakārtoti teikumi*). They point out that the clauses in compound sentences are of equal status. Halliday (2004, 373–374) refers to such a status of the clause relation as parataxis, whereas Biber et al. (1999, 136), Blinkena (2009, 157), Bušs et al. (2007, 335), and Downing, Locke (2009, 26) as coordination (in the Latvian language: *sakārtojums*).

The second group of sentences that contain more than one clause are labelled complex sentences (in the Latvian language: *salikti pakārtoti teikumi*) by linguists such as Auziņa et al. (2013, 832), Downing, Locke (2006, 272), and Yule (2006, 266). They explain that the clauses in these sentences are of unequal status and such clause relation is termed hypotaxis by Halliday (2004, 373–374) or subordination (in the Latvian language: *pakārtojums*) by Biber et al. (1999, 136), Blinkena (2009, 157), Downing, Locke (2009, 26), as well as Quirk et al. (1985, 44).

Apart from compound and complex sentences, the so-called compound-complex sentences (in the Latvian language: *jaukti salikti teikumi*) are distinguished by the researchers of the Latvian language. They contain several clauses of equal as well as unequal status (e.g. see Auziņa et al., 2013, 831). In this connection Downing, Locke (2006, 273–174) explain that coordination and subordination of clauses in a connected discourse ‘may be [...] complex and variable’, so they expand the notion of a complex sentence by claiming that ‘a complex sentence can consist of any number of clauses of different types and in different combinations’.

Complex sentences or sentences with unequal relation of clauses are further described by linguists from two angles – the type of dependent clause and the semantic degree of subordination in its relation to the independent clause, both of which play an important role in punctuation solutions in English. According to traditional grammar, there are three main types of dependent clauses that are named after the relevant parts of speech: Noun, Adjective and Adverb clauses; however, there are terminological variations in the description of Noun and Adjective clauses. Noun clauses (Yule, 2006, 160) are referred to as nominal clauses by Biber et al. (1999, 193) or complement clauses by Trask (1993, 10) and Hurford (1994, 232), for example. The term ‘adjective clause’ or ‘adjectival clause’ seems to have been abandoned by linguists in favour of the term ‘relative clause’ (Biber et al., 1999; Downing, Locke, 2006; Yule, 2006, etc.). The term ‘adverbial clauses’ is generally accepted and has been widely used by linguists.

Although the previously mentioned dependent clauses are in a subordinate relation to the relevant independent clauses, the degree of subordination of dependent clauses

can vary in a complex sentence. Downing, Locke (2009, 275) emphasise that it is important to '[...] adopt the view that dependency is not an absolute property, but rather a question of degree'. They explain that the dependency degree shows the degree of integration of a dependent clause as 'perceived or imagined by the speaker or writer between events'. The tightest integration is referred to as embedding by linguists, which means that the dependent clause functions as a constituent of the independent clause.

On the basis of the aforesaid classification of complex sentences, some linguists point out that the English dependent clauses, which are embedded within complex sentences (e.g. Downing and Locke, 2009) and, therefore, do not require separation by a comma, are: nominal *that-clauses* (Example 1); nominal *wh-interrogative clauses* (Example 2) including *whether* and *if*, as well as relative (restrictive) *that* and *wh-clauses* (Examples 3 and 4).

- (1) He saw **that** the bottles were empty.
- (2) I'll ask **where** the nearest underground station is.
- (3) The house **that** they live in is large.
- (4) Perhaps the people **who** were waiting are still there.

The perception of the degree of dependent clause subordination plays an important role in complex sentence punctuation in English. As it is explained in English punctuation manuals, a comma is used 'to make a slight break' (Davidson, 2005, 78), and for this function Lukeman (2007, 31) has metaphorically described a comma as 'the speed bump of the punctuation world' because 'with its power to pause, the comma controls the ebb and flow of a sentence'. This implies that the pause marked by a comma is relevant in the cases where the dependent clause is less tightly integrated. Comma variation caused by semantic dependency in English complex sentences obviously explains why English punctuation is considered a challenging issue of text composition by Latvian students of English (Farneste, 2006a, 2006b).

The interrelation of an embedded subordinate clause separated by a comma reveals a noticeable, though neglected, aspect of comma use differences in Latvian and English complex sentences. Although some researchers discuss the semantic relation of clauses in Latvian complex sentences (e.g. Beitiņa, 2009), this phenomenon does not have the same impact on the punctuation in Latvian texts as it has in texts written in English. In Latvian complex sentences, as Blinkena (2009, 176) states, all dependent clauses, irrespective of their semantic dependency degree, are normally separated by a comma (see Example 5 and its translation into the English language – Example 6).

- (5) Jāatzīst, **ka** lielie burti atvieglo lasīšanu.
- (6) It has to be noted **that** capital letters relieve reading.

Thus, the theoretical analysis has revealed that comma use/non-use in English complex sentences, containing an independent clause followed by a dependent clause, is based on the type (e.g. nominal, relative) and semantic degree of subordination. The neglect of semantic dependency might cause punctuation confusion as, for example, in

the following case: a nominal *who-interrogative clause* (Example 7), which is embedded in the main clause and, therefore, is not separated by a comma, might be confused with a non-embedded (i.e. *non-restrictive/non-defining*) relative clause introduced by *who* (see example 8) and, therefore, separated by a comma. The loose dependency degree of all the *wh-non-restrictive/non-defining* relative clauses should also be marked by the comma use.

(7) They know **who** has left this paper on the table.

(8) I'll give the CD to Ben, **who** likes music.

Method, procedure and results

The material for analysis was extracted from a Latvian student-composed text corpus and the analysed set contained 104 texts of approximately 21,319 words. The texts were written by students before their participation in a writing course. Their selection aimed at revealing if the students had used the above-discussed clause types and patterns, namely, the embedded clauses (nominal and relative) that are preceded by an independent clause, and what punctuation solutions they offered to serve as a departure point for the development of their writing skills.

In order to extract these clause types and patterns, corpus-based data extraction procedures were applied (McEnery et al., 2006; McEnery, Hardie, 2012). The clauses were extracted with the help of the *AntConc* concordancer. Seeking to ease data extraction, the texts were part-of-speech tagged by CLAWS7, which contains tags for the linguistic features functioning as subordinating conjunctions in complex sentences that are in the focus of this study (see Table 1).

Table 1. Tags

Tags	Linguistic features	Subordinate clauses
CST	<i>that</i> as conjunction	Nominal <i>that-clauses</i> Relative <i>that</i> clauses
VV* * RRQ/CSW/DDQ	<i>wh-</i> general adverb (where, when, why, how)	Nominal <i>wh-interrogative</i> clauses
PNQS	Subjective <i>wh-pronoun</i> (who)	Relative <i>who-restrictive</i> clauses
PNQO	Objective <i>wh- pronoun</i> (whom)	Nominal <i>whom-interrogative</i> clauses Relative <i>whom-restrictive</i> relative clauses

The tag CST of CLAWS7 was used as a query to extract nominal and relative *that-clauses* as conjunction concordance lines that were sorted into nominal and relative

that-clauses. The obtained concordance lines were sorted to distinguish both clause types.

The students' texts, as expected, contained both nominal and relative *that*-clauses. The majority of these clauses (Table 2) were correctly presented by the students as embedded, as they had not been separated by a comma. Example 8 shows an embedded nominal clause, and Example 9 an embedded relative clause that, due to their strong semantic relation, are not separated by a comma. The students embedded their nominal clauses in such common structures as *I think; I would like to tell; people think; I knew that; I must say; I can conclude*' etc., as seen in Examples 8 and 10.

(8) Some people think **that** travelling is a bad and useless way of spending money.

(9) There are many routes **that** go through the city.

There were also occasional dependent nominal and relative *that*-clauses that, irrespective of their embedded information, were separated by a comma. The two sentences in Example 10 show comma use in nominal clauses, but example 11 shows comma separation of a restrictive relative clause.

(10) I believe, **that** this way of transport is quite developed. // I think, **that** it is a beautiful city.

(11) It has been circumstances, **that** are the most negative factor.

The overall percentage of nominal and relative clauses that were/were not separated by a comma was equal. It has to be noted that commas would normally be used in such sentences in the Latvian language, so the redundant comma use might have occurred due to the impact of their use in the Latvian language.

Table 2. Dependent clauses in student-composed texts

Dependent clause types	Number of dependent clause examples:	
	without a comma separation	with a comma separation
Nominal <i>that</i> -clauses	89 (88%)	12 (12%)
Relative <i>that restrictive</i> clauses	61 (88%)	7 (12%)
Nominal <i>wh-interrogative</i> clauses	11 (78%)	3 (22%)
Relative <i>who-restrictive</i> clauses	57 (85%)	10 (15 %)
Nominal <i>whom-interrogative</i> clauses Relative <i>whom-restrictive</i> relative clauses	0	0

The extraction of nominal interrogative clauses required the application of several tags. The tag RRQ extracted *wh-general adverbs* and the output included not only *nominal wh-interrogative* clauses, but also e.g. direct questions and adverbial clauses. Therefore, in order to extract nominal clauses that stand after an independent clause, the search was narrowed to VV* * RRQ and other tag modifications (Table 1). They were all placed

in the advanced search option, as well as the output concordance lines were examined and sorted to check each instance for relevance. Presuming that the search options had enabled the extraction of all nominal interrogative clauses found in the texts, the results showed that students had used a small number of them. Although the majority of the clauses were not separated by a comma (as in Example 12), there were also cases where the students had separated these clauses by a comma, as in Example 13, where the author might have intended to insert a pause after ‘wonder’ for a dramatic effect. Obviously a dash would have been more relevant in that case.

(12) I didn't know **where** they were.

(13) One might wonder, **what** will happen.

It has to be noted that the number of these clauses extracted is too modest to serve for generalisations.

Although the PNQS tag extracted *who* in different position within a sentence, e.g. at the beginning of sentences, which were just a few cases in the students' texts, it allowed convenient extraction of relative *who*-clauses, which in their turn could be sorted into restrictive and non-restrictive clauses to check the relevance of comma use. Relative *who*-restrictive/embedded clauses that are normally used without a comma can cause punctuation challenges, as *who* is used to introduce both restrictive /embedded and non-restrictive/non-embedded clauses. Indeed, *who* relative restrictive clauses turned out to be challenging. Apart from numerous relevantly punctuated instances (Examples 14), there were also cases of the irrelevant insertion of commas (Examples 15).

(14) Of course there are people **who** mostly prefer walking.

There are a lot of people **who** are driving to work by car.

People **who** do not have their own car.

(15) There is anybody, **who** would love.

I feel it more than people, **who** live in a city.

(16) Now factories **who** make cars think how to make their cars friendly to nature.

In addition, as can be seen in Example 16, there are clauses in which the use of *who* is doubtful, as *who* wrongly refers to *factories*, though it normally relates to humans. It is also noteworthy that cases of non-restrictive *who* relative clauses were not found in students' texts, which means that they did not display varied approaches to sentence structure to reach the communicative purpose of the text.

The PNQ0 tag was employed to check if the cases of nominal and relative *whom* clauses were found in the texts. It was determined that these clauses had not been used by the students.

Conclusions

1. The comparative analysis of the theories has revealed that there are variations in the terms that linguists use to describe sentences containing more than one clause.

2. The analysis of the results of student-composed sentences, on the one hand, reveals their awareness of the tight semantic interrelation of such clauses; yet, on the other hand, it also reveals cases of comma overuse in the complex sentences that contain either relative or nominal clauses. It also uncovers that the students might have proposed more varied nominal and relative clause options in their texts.

3. The students' own created examples can serve for an in-depth discussion of the degree of dependent clause integration in particular communicative situations, which might lead to the composition of more varied English complex sentences to suit the communicative purpose of their texts.

4. These results also call for a more profound focus of the students' attention on the nuances of clause dependency in English by comparing English and Latvian sentence structures, as well as the degree of their semantic relation and punctuation.

Further research into various clause types in student-composed texts of different communicative purposes would be significant to reveal the issues related to the syntactic complexity of their texts.

References

- Auziņa, I., Breņķe, I., Grigorjevs, J., Indričāne, I., Ivulāne, B., Kalnača, A., Lauze, L., Lokmane, I., Markus, D., Nītiņa, D., Smiltņiece, G., Valkovska, B., Vulāne, A. (2013). *Latviešu valodas gramatika*. Rīga: Latviešu valodas unstitūts.
- Beitiņa, M. (2009). *Mūsdienų latviešu literārās valodas sintakse*. Liepāja: Liepājas Universitāte.
- Biber, D., Johansson, S., Leech, G., Conrad, S., Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. Longman.
- Blinkena, A. (2009). *Latviešu interpunkcija*. Rīga: Zvaigzne ABC.
- Bušs, O., Joma, D., Kalnača, A., Lokmane, I., Markus, D., Pūtele, I., Skujiņa, V. (2007). *Valodniecības pamatterminu skaidrojošā vārdnīca*. LU Latviešu valodas institūts.
- Ceplītis, L., Rozenbergs, J., Valdmanis, J. (1989). *Latviešu valodas sintakse*. Rīga: Zvaigzne.
- Davidson, G. (2005). *How to Punctuate*. Penguin Books.
- Downing, A., Locke, P. (2006). *English Grammar*. Routledge: London and New York.
- Farneste, M. (2006a). *Punctuation as a Means of Text Cohesion in Written Communication*. Proceedings of the Fourth Riga International Symposium *Pragmatic Aspects of Translation*, Riga: University of Latvia, 36-48.

- Farneste, M. (2006b). Comparative Analysis of Comma Usage: English vs. Latvian *Valoda dažādu kultūru kontekstā*. Daugavpils Universitāte. Humanitārā fakultāte. Daugavpils: Saule, 372–379.
- Halliday, M. A. K. (2004). *An Introduction to Functional Grammar*. Hodder Arnold.
- Hurford, J. R. (1994). *Grammar: A Student's Guide*. Cambridge: Cambridge University Press.
- Lukeman, N. (2007). *The Art of Punctuation*. Oxford: Oxford University Press.
- Lyons, J. (1995). *Linguistic Semantics: An Introduction*. Cambridge: Cambridge University Press.
<http://dx.doi.org/10.1017/CBO9780511810213>
- McEnery, T., Xiao, R., Tono, Y. (2006). *Corpus-based Language Studies*. Routledge.
- McEnery, T., Hardie, A. (2012). *Corpus Linguistics*. Cambridge: Cambridge University Press.
- Trask, R. L. (1993). *A Dictionary of Grammatical Terms in Linguistics*. London: Routledge.
- Yule, G. (2006). *Oxford Practice Grammar*. Oxford: Oxford University Press.
- Quirk, R., Greenbaum, R., Leech, G., Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. Harlow: Longman.

Tools

AntConc available at: <http://www.laurenceanthony.net/software.html>

CLAWS7 Tagger available at: <http://ucrel.lancs.ac.uk/claws/trial.html>

Sudėtiniai sakiniai ir jų skyryba angliškai parašytuose Latvijos studentų tekstuose

Zigrīda Vincēla

Latvijas universitēto Humanitāriņū mokslū fakultetas, Visvalža g. 4 A, LV-1050, Latvija,
zigrida.vincela@lu.lv

Santrauka

Latvių ir anglų kalbų sudėtinuose sakiniuose gerokai skiriasi kablelio vartojimas, todėl latviams, rašantiems angliškai, iškyla sunkumų. Iš esmės prijungiamasis dėmuo, kuris eina po pagrindinio dėmens, latvių kalboje atskiriamas kableliu (Blinkena, 2009), o anglų kalboje kablelio rašymas ar nerašymas priklauso nuo semantinio pagrindinio ir prijungiamojo dėmens ryšio (Downing, Locke, 2006). Šie kablelio vartojimo skirtumai, aptarti kitų tyrėjų (Farneste, 2006b), gali kelti sunkumų latviams studentams rašant ir skiriant sudėtinius sakinius anglų kalba.

Šio tyrimo tikslas – išsiaiškinti, kaip latvių studentai skiria sudėtinius sakinius anglų kalba parašytuose tekstuose. Medžiaga tyrimui imama iš studentų rašto darbų duomenų bazės. Šiame tyrime dėmesys skiriamas šalutiniams veiksnio, papildinio ir pažyminio sakiniams, kurie semantiškai itin glaudžiai susiję su pagrindiniu dėmeniu. Paprastai, kaip nurodo Downingas ir Locke'as (2006), tokie prijungiamieji sakiniai anglų kalboje nėra atskiriami kableliu. Tyrimas rodo, kad glaudus semantinis dėmenų ryšys studentų dažniausiai atpažįstamas ir skiriamas pagal taisykles. Vis dėlto pastebima ir perteklinės skyrybos atvejų. Šie rezultatai skatina atkreipti studentų dėmesį į semantinius anglų kalbos sudėtinių sakinių dėmenų ryšius ir jų įtaką skyrybai bei išryškinti latvių ir anglų kalbų skyrybos skirtumus. Be abejo, ateityje būtina ir nuodugnesnė studentų tekstuose vartojamų įvairių sakinių tipų ir jų skyrybos analizė.

Esminiai žodžiai: *sudėtinis sakiny, dėmuo, skyryba, semantinė priklausomybė.*

Gauta 2015 01 13 / Received 13 01 2015
Priimta 2016 02 01 / Accepted 01 02 2016