

Industrial impact on topics and types of Master's theses Empirical study of software engineering theses made in 1990-2016.

H.-M. Järvinen

professor

Tampere University of Technology

Tampere, Finland

E-mail: hannu-matti.jarvinen@tut.fi

T. Mikkonen

professor

Tampere University of Technology

Tampere, Finland

E-mail: tommi.mikkonen@tut.fi

Conference Key Areas: University-Business cooperation, Sustainability and Engineering Education

Keywords: software engineering education, thesis, industrial impact

EXTENDED ABSTRACT

INTRODUCTION

One of the ways universities and industry co-operate is making the master's theses on the topics of industrial partners. In this paper 578 theses on software engineering from 1990 until 2016 are evaluated to see how the needs of the industry on information technology in Finland have affected the topics, type, language and orientation of the theses. Also the size of the company and the gender of students were recorded as well. All the theses have been supervised by either of the authors and they represent about 30 percent of theses on software engineering at Tampere University of Technology.

Our strongest hypothesis was that during 2000-2005 golden era of Nokia would affect greatly on the numbers so that the major part of the theses were made for a large company, mobility is one of the most general topics and there are several constructive theses that are part of bigger projects. Other initial hypotheses were that the number of theses in English has been increased since 1990, the number of females has been the same or increasing slightly, and the orientation of the theses

(constructive or research-oriented) has not changed much, the constructive ones being much more common.

The results partly proved the hypotheses, but interestingly enough, we got some surprises especially on the language of the theses and the gender on students.

1 METHODOLOGY

The theses were classified in six dimensions: the sponsor, the type, the language, the orientation, the topic, and gender of the author of the thesis. The classification was made by the authors of the paper for the theses they had supervised; the official grading statements of the theses and the theses themselves were used as the basis of the classifications.

The sponsor had six possible classes: large (hundreds of employees), medium-size, or small company (under 40 employees), the university of the authors, other academic or research institution, or the student's own topic.

The type of the theses was classified to be either an independent study, independent part of a bigger project, part of a project, or an improvement of a (software) process. The orientation was either constructive, i.e., the goal was to solve a practical problem, or research-oriented. Since many of the theses have some characteristics of both of the orientations, the dominating one was selected.

The default language of the university is Finnish, but other languages are accepted, too. In practise, only English is used in addition to Finnish as the language of theses.

The topic of the theses was classified to belong on of the following classes: information systems, traditional computer science, mobile systems, embedded or operating system related, or web or user interfaces. The gender was either male or female.

2 HISTORICAL OUTLINE

There have been several points in economic or educational surroundings that most probably have affected the topics of theses.

The curriculum of information technology in TUT was created in 1985, so with the average of completion time of 5.5-6.6 years, the first theses in this study are either early graduates of that programme, or students of electrical engineering. However, the real rush of students started in 1995 when the intake of the curriculum was greatly increased. This caused the yearly number of the theses to increase after 2000.

In 1992, deep depression in Finland took place because of the devaluation of Finnish mark. On the other hand, Nokia was becoming more and more important company in the Finnish economy. This affected especially information technology students and theses. Since Nokia used many subcontractors, IT industry in general was growing fast. Nokia was at its high in 2000-2007 and dominated IT in Finland. In 2008, the first downsizing of Nokia started together with the general depression of economy. Until 2010, Nokia was developing its own operating system for cellular phones; this can be seen in the topics section later on.

In late 1990's, IT was one of the most famous curricula in Finland and there were no problems getting students in. In 2002, the so called "IT bubble" took place, and interest in IT studies decreased significantly. After that, the intake has been

decreased to be less than half of what it used to be at its best (currently 90, used to be even 220).

Five groups were formed according to the year when the theses were completed: 1990-1999, 2000-2004, 2005-2009, 2010-2014, and 2015 until April 2016. During the first ten years, only the first author supervised the theses (not every year) and it contains only 37 theses. Period 2000-2004 contains 129 theses, period 2005-2009 161 theses, period 2010-2014 203 theses and the last period from 2015 until April 2016 51 theses. Year 2010 is exceptional in the number of theses, because of the reorganisation of the curriculum, it was the last possible year for many students to graduate using their original study plan. Hence, 67 students graduated that year. On the other hand, the year 2008 and depression can be clearly seen in the number of graduates: only 20 theses.

The software process was based on the so called waterfall model for years. This model is considered outdated and is being replaced by agile methods, especially in last decade. This has clearly caused a great increase on theses having process improvement as their topic.

3 FINDINGS

3.1 The sponsor of the theses

The sponsors of the theses are depicted in Figure 1. As was anticipated, during 2000-2004 the majority of the theses (almost 60%) were made for large companies, as Nokia dominated in the Finnish economy. After 2005 the share of large companies has been decreasing to the current level of 10 percent. The winners have been medium-sized and small companies, but also the student's own topic.

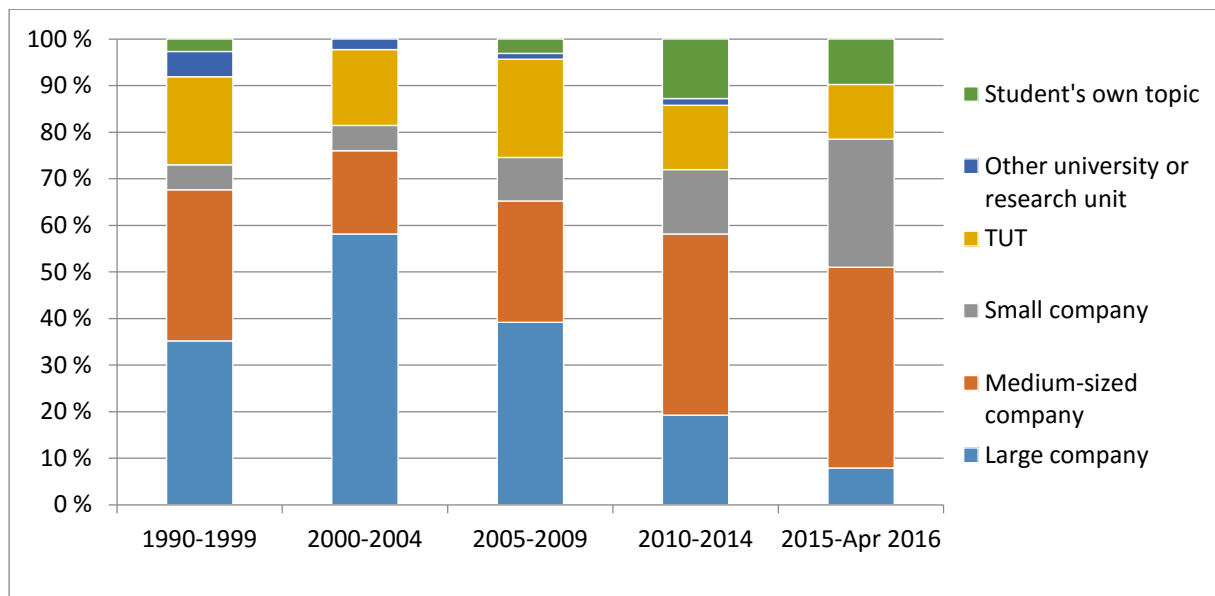


Figure 1 Sponsor of the theses

The downsizing of Nokia explains mostly why the share of large companies has been decreasing. Since Nokia was not the only one large company offering topics for theses, the current small share of the large companies cannot fully be explained by Nokia. One possible explanation is that since most of the large companies in Finland produce machinery for other enterprises, the depression has decreased their export

and hence also their ability to hire new employees and interns. In the other hand, Nokia's downsizing created new companies with personnel that had earlier experiences on offering master theses. This partly explains why so many theses are currently made for small and medium sized companies.

An interesting finding is that the number of theses having a topic suggested by the students themselves has increased to 10 percent, even above. This may reflect problems in getting suitable topics because of the depression of economy, but a new phenomenon is that the students have not asked for topics from the university as they have done earlier. This can be seen if compared the years 2005-2009 with the years 2010-2014 in Figure 1.

3.2 The type of the theses

The types of the theses are illustrated in Figure 2. Theses classified as part of a bigger project are virtually vanished by now; they are mostly made in large companies. Most of the theses are either independent studies where the student has been the only one working on the topic, or independent part of a bigger project, where there has been clear subdomain given to the student. Topics related to improving the software process have increased substantially. One of the reasons is switching for agile methods, which has created new kinds of ways to organise working. Also the increasing share of medium-sized companies is a partial explanation of this, since just after the start up goes up, it is relative easy to handle, but a growing company needs new ways to answer its organisational problems. Since agile methods are still young and there are a number of software engineering tools where to select form, special attention has to be paid how to apply them when the company is growing.

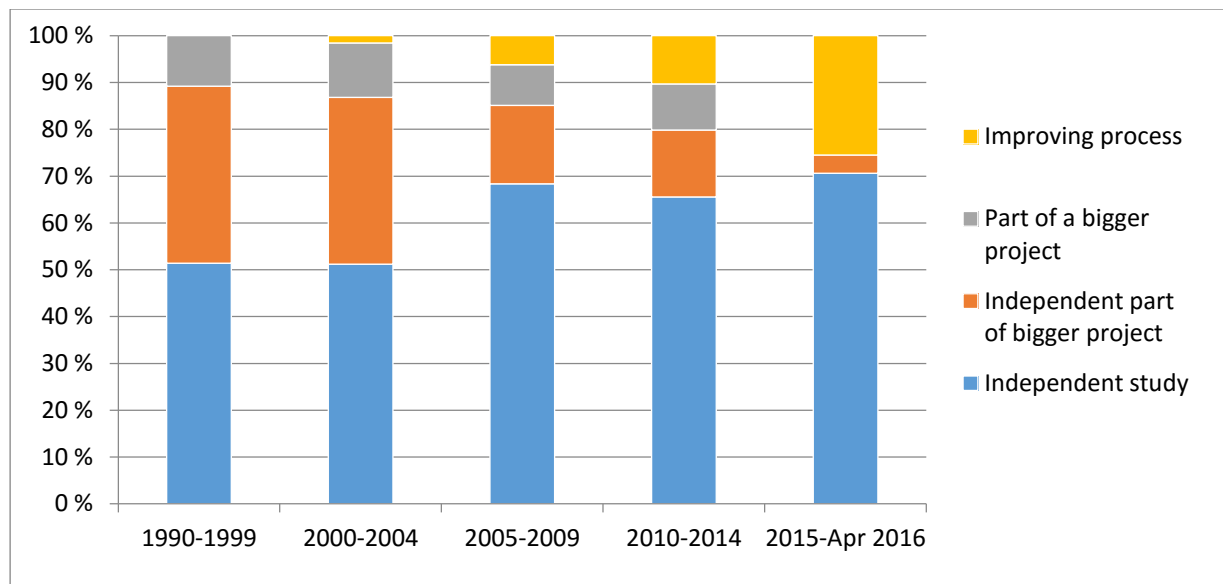


Figure 2 Type of the theses

3.3 The orientation of the theses

Most of the master's theses in Finland are made for enterprises. Hence, it is not a surprise that most of the theses are also constructive in nature. Constructive theses contain practical problems to be solved; often these include the utilisation of some new technology or selection of the new generation technology. Somewhat surprisingly, many of the theses funded by universities are constructive, too. The

reason is that doctoral students work normally on research problems and master's theses are more practical work intended to support the research. The orientation of the theses is shown in Figure 3.

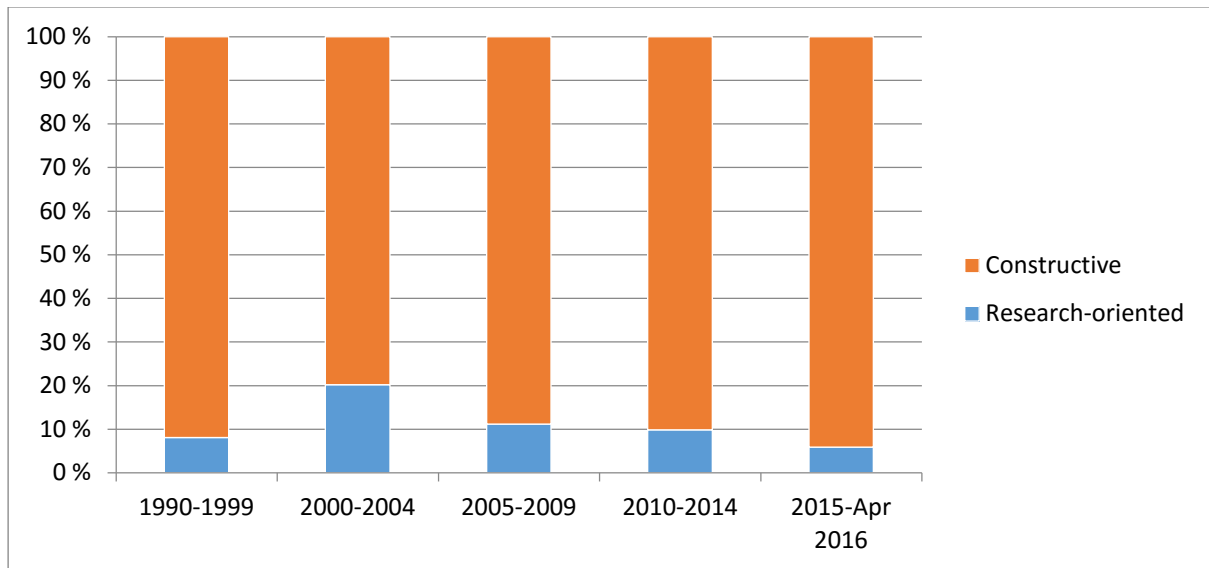


Figure 3 Orientation of the theses

3.4 The language of theses

During the first decade (1990-1999), only 10 percent of the theses were written in English. Contrast to the years 2000-2004 is huge since within this period about 50 percent of the theses were written in English. After that the share of English has decreased to one third of the theses (see Figure 4).



Figure 4 Language of the theses.

The decrease in using English as the language of theses is explained by the decreasing number of theses made for large companies, since most of the small and medium-sized companies favour Finnish in their theses (about 83 % and 77 %, respectively. The large share of theses in English during 2000-2004 is explained by English being the internal language of Nokia.

However, these findings are somewhat in contradiction to our assumptions. We had a feeling that the number of theses in English had gradually increased during the

years, especially because the international master's programme of the university was extended to cover computer science, too.

3.5 The topic of the theses

Figure 5 contains the distribution of the topics of theses. The number of theses on information systems, i.e., theses that concentrate on how to use information technology to assist either commercial or other processes of stakeholders, used to be quite small until 2010, after that its share has increased to a little bit over 10 percent.

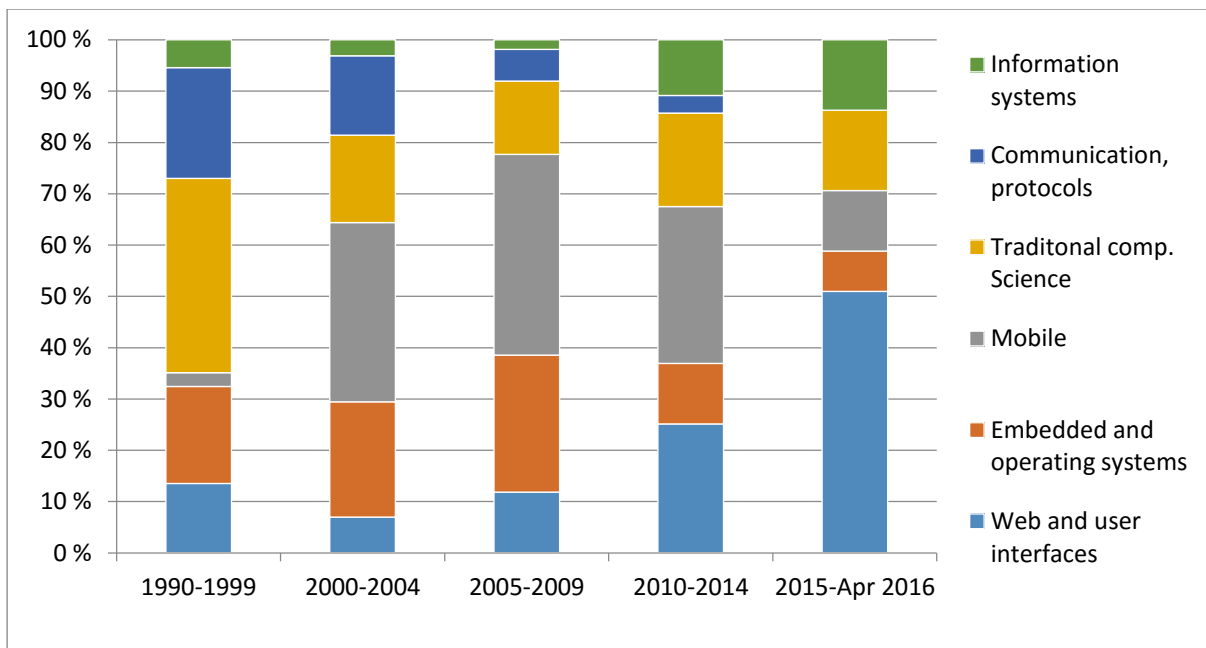


Figure 5 Topic of the theses.

Traditional computer science has had a relative steady share of the topics through the years except the first decade when it dominated. Work related to embedded systems and operating systems used to have a steady share too, but this share has recently dropped, perhaps due to closing down of the Symbian evolution. As anticipated, mobile systems were on their top from 2000 to 2009, but have shrunk after that. Web-related topics have had a significant increase in their share in last few years, now covering over 50 percent of the theses.

3.6 The gender of the authors

In technology areas, the majority of the students has been male. For some reason, this seems to be even stronger phenomenon in information technology. The authors had an impression that the situation has been quite steady for years. Surprisingly, the number of theses written by female students first decreased from 11 to 3 percent, then increased again to 14 percent, as can be seen from Figure 6. There is a partial explanation to this. During 1990-1999, the first author supervised usability-related theses, until the faculty got a professor for usability. Since many female students select usability as their major subject, this may explain the decrease, but not the increase.

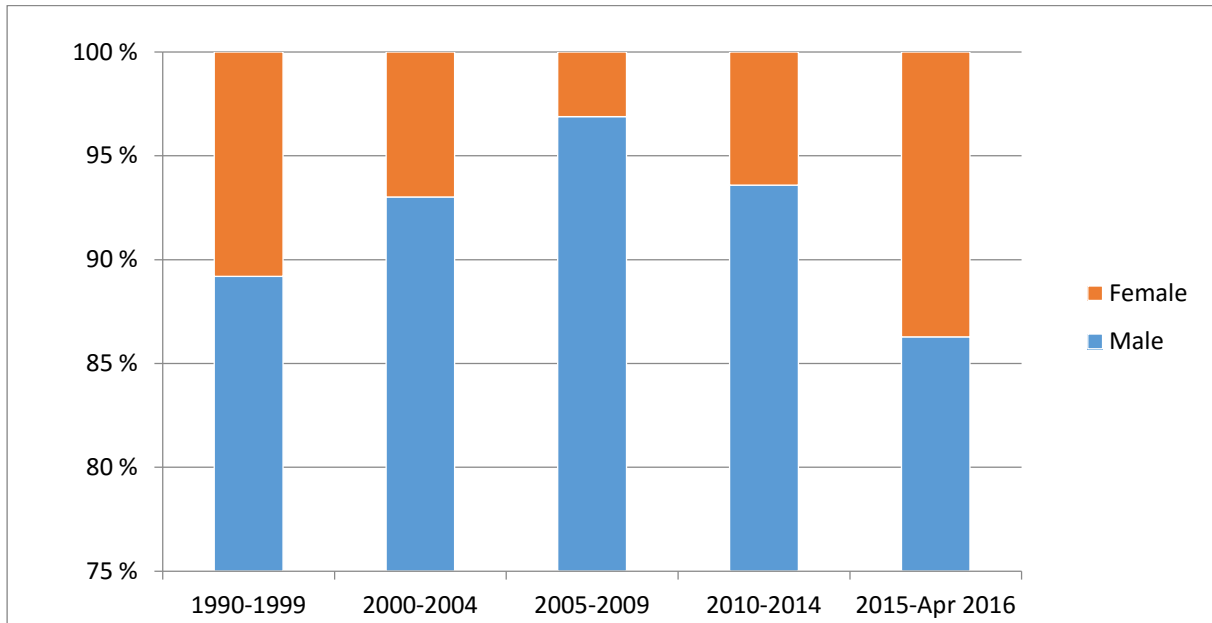


Figure 6 Gender of the author of the theses (note the scale of Y-axis)

However, the gender-related variables have the problem that there are only 38 females in the overall study, so one has to be careful when interpreting the numbers.

4 RELATION OF INSPECTED VARIABLES

In this section, some of the relations between the variables of the study are highlighted. These are given as facts without deeper discussion and their purpose is to make it easier to understand the timelines discussed above.

4.1 Sponsor-related notes

In Figure 7, the relative shares of constructive and research-oriented theses are illustrated. Clearly universities and research units offer more research-oriented topics, but a little bit surprisingly, students' own ideas are on about the same level with these. The small companies offer mostly constructive topics. On the other hand, the small companies offer theses that are independent studies, only the "own ideas" of students overcome these (see Figure 8).

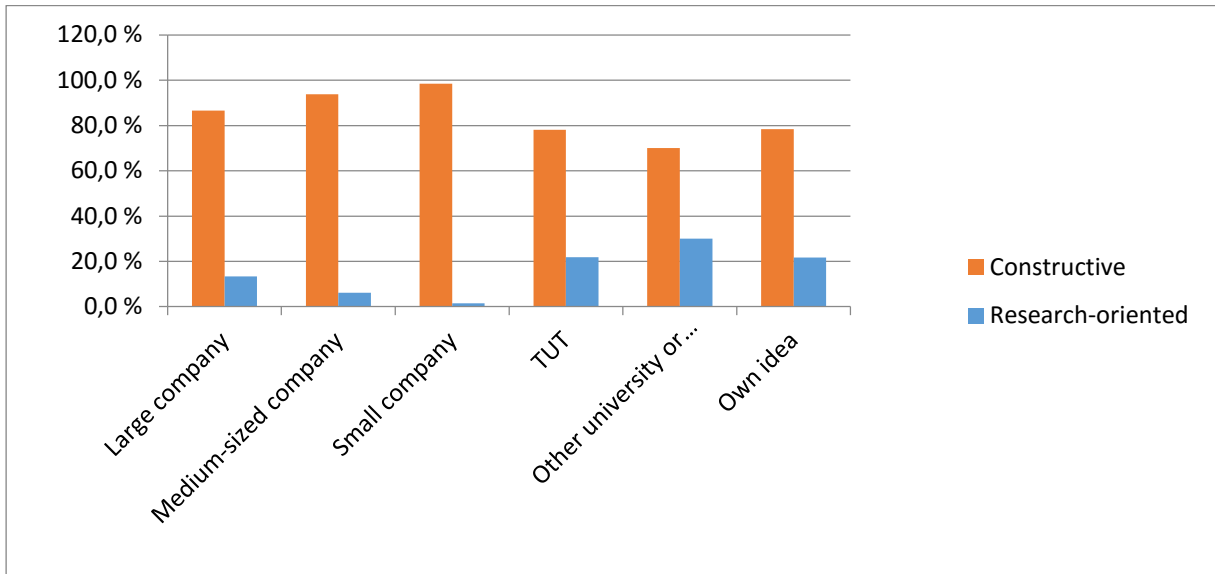


Figure 7 Orientation of theses by the sponsor

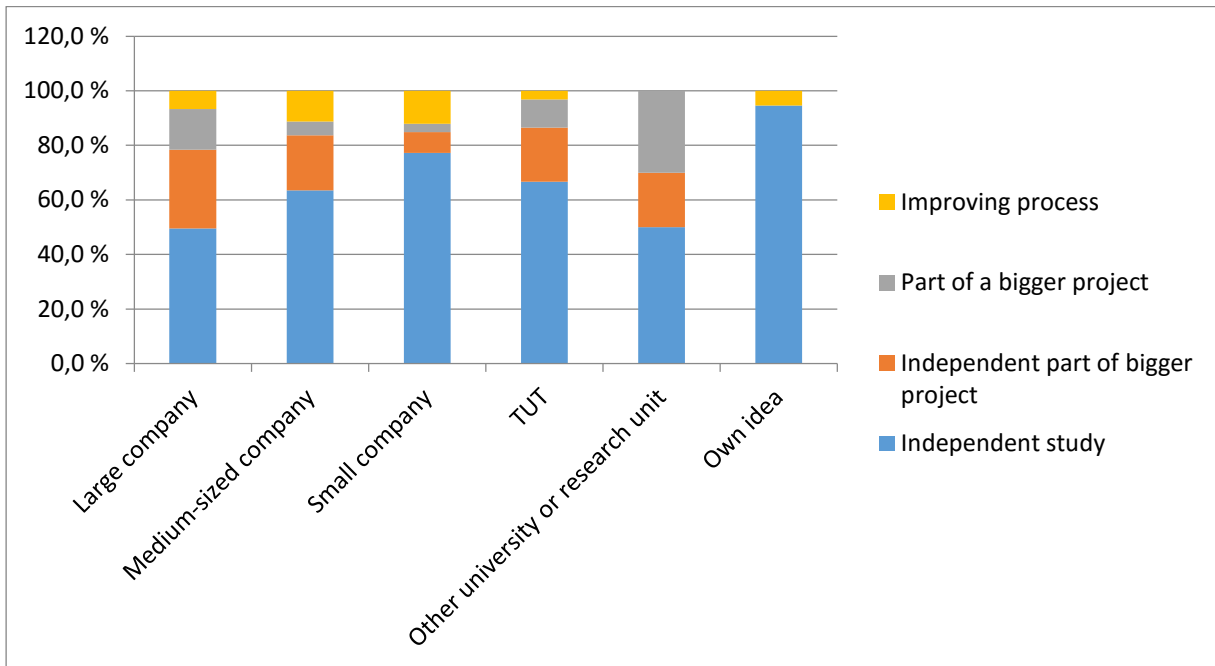


Figure 8 Types of theses organised by sponsor

4.2 Language-related notes

As noted in Section 3.4, small and middle-sized companies favour Finnish as the language of the theses. There are surprisingly high differences in the language between topics, illustrated in Figure 9 Theses written in Finnish or English by the topic.

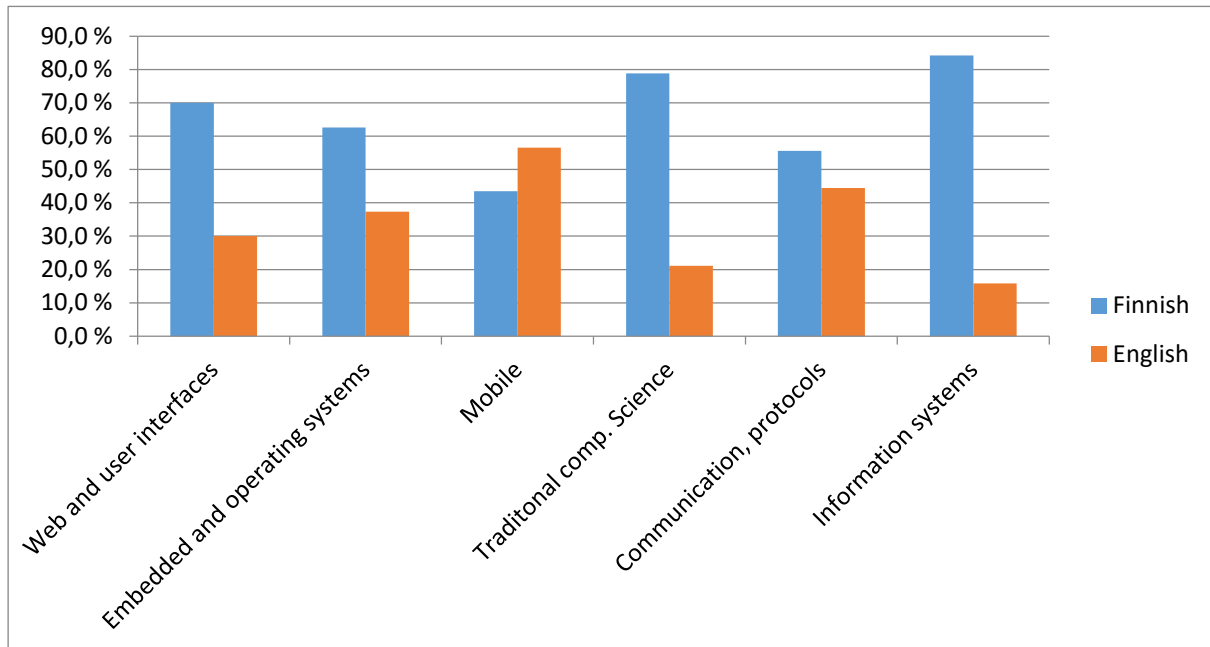


Figure 9 Theses written in Finnish or English by the topic.

On the other hand, it was expected that research-oriented theses are more common in English than the constructive one, as depicted in Figure 10 Language of the theses by orientation. Figure 10.

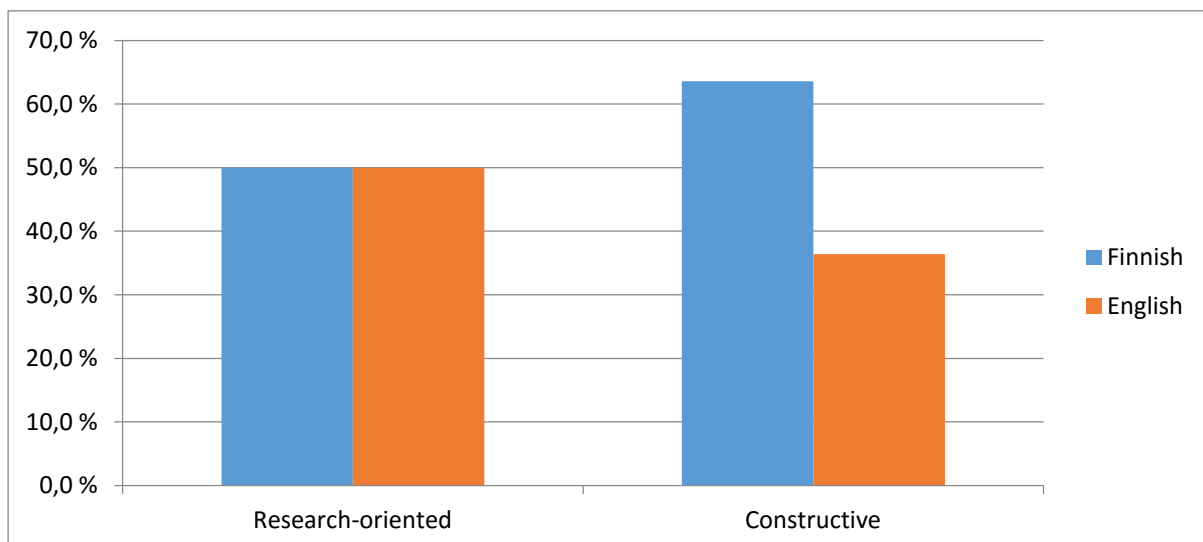


Figure 10 Language of the theses by orientation.

4.3 Gender-related notes

The gender has almost non-existent impact on the language, sponsor, or orientation of the theses. However, there is a clear impact on the topics and type of the theses,

as can be seen from Figure 11 and Figure 12. The biggest difference found in Figure 11 is in the number of students working on improving a process.

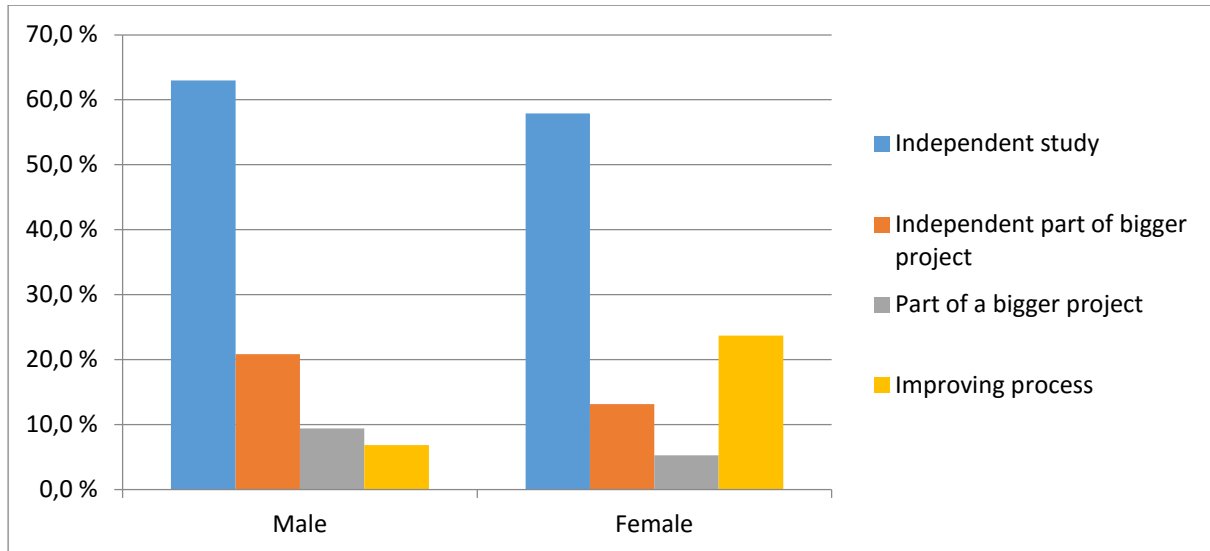


Figure 11 Percentage of male and females to select different types of theses.

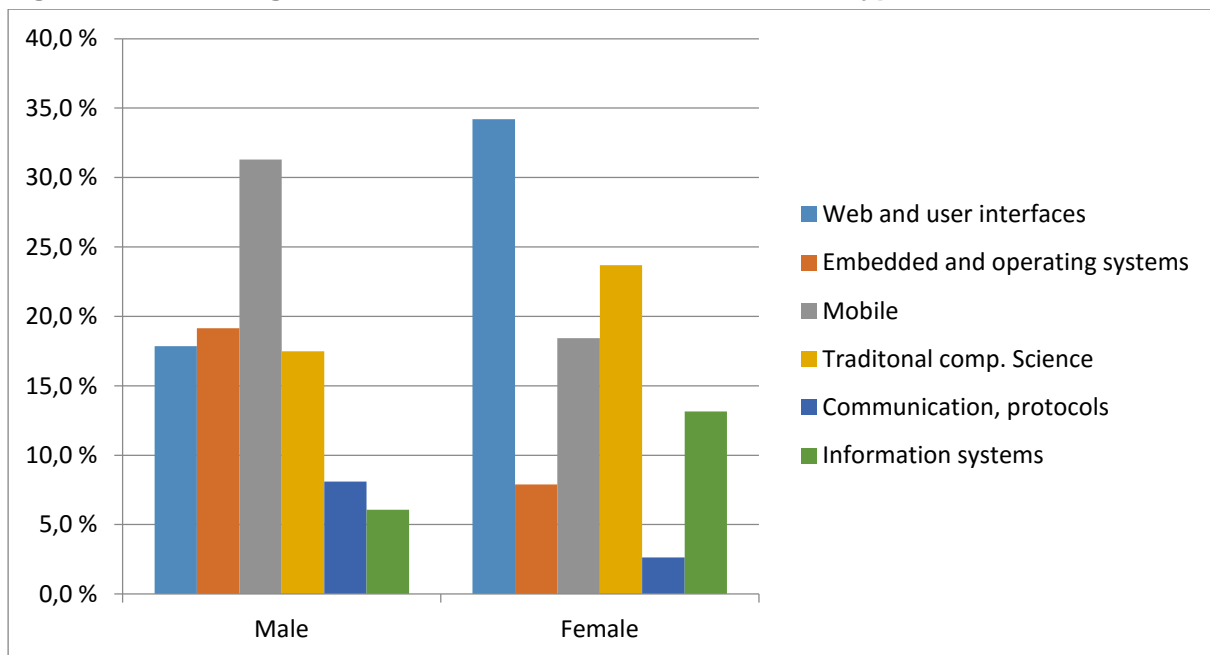


Figure 12 Percentage of male and females to select different topics of theses.

In Figure 12, there are several variables that clearly differ. Females favour web and user interfaces, information systems, and even traditional computer science. On the other hand, a relative big share of males selects embedded and operating systems, and communication protocols.

As noted above, due the small number of females in the overall study, one has to be careful when interpreting the numbers.

5 SUMMARY

There have been clear changes in sponsors, topics and types of the theses during the years; many of them can be explained by the changes in economy and technology. Changes in the language exist, but they seem to be stabilized. In

general, theses which are part of a research project or product development are more often written in English than theses aiming to improve the internal processes of the company.

The relative share of research-oriented and constructive theses has been quite stable during the years. The share of genders was changing, but there is not evident reason to it. Since the number of females is low in comparison with the total number, small absolute changes may exaggerate the changes.

Some relations between variables were included here, mostly to make it easier to interpret the results in Section 3. In the future, they should be studied in more detailed to find out the phenomena behind them.