# Preface

This dissertation is all about the fascinating topic of syntax. Such a statement will probably be frowned upon by many computer scientists and programmers. Syntax, in particular textual syntax, seems to be a rather unpopular topic these days. Grammars and parsing are not particularly active research topics, resulting in a lack of innovation in one of the most foundational topics of computer science. In fact, textual syntax is so unpopular that people try to avoid it altogether and attempt to move to development methods without a textual syntax. However, instead of abandoning textual syntax for its problems, I suggest we solve those problems and advance the state of the art in grammars and parsing.

About six years ago I started working with the infrastructure for parsing programming languages in the program transformation system Stratego/XT. Stratego/XT employs the grammar formalism SDF and the scannerless generalized LR parser SGLR, which is the target of the SDF parser generator. The combination of SDF and SGLR is very easy to use. Indeed, I have been using the SGLR parser for years without knowing at all how it works! The essential advantages of scannerless parsing over parsing with a separate scanner are *(1)* lexical disambiguation by context, *(2)* full description of the syntax of a language in a uniform grammar formalism, and *(3)* expressive lexical syntax. These advantages are easy to understand. They were originally motivated by application to minor issues in parsing existing programming languages, such as the lexical ambiguity between subrange types versus floating point literals in Pascal. Although these examples clearly illustrate the advantage of scannerless parsers, the problems with implementing parsers for these languages are not serious enough to make a strong case for scannerless parsing.

The first time Eelco Visser presented his then latest ideas on the use of concrete object syntax [Visser 2002], where the syntax of an object language is embedded in a metalanguage, I was rather skeptical about combining languages in this way. Slowly, however, I realized the full potential of scannerless generalized parsing for combining languages. This thesis is largely about explaining this potential, contributing technical bits and pieces here and there. Indeed, I consider my work to be an exploration in what we could achieve once we are liberated from the limitations of conventional parsing techniques. This explains the title, which is a reference to the beautiful song *Exercises in Free Love* by Freddie Mercury[1] (various puns intended). The Dutch title is a pun as well, referring to the almost military operations I usually set up to finish our papers before a conference deadline.

There is still a lot of work left to do. Reflecting on my work, I feel that one of the main contributions is that I provide a strong motivation for rethinking

[1] Exercises in Free Love is a song performed without lyrics as a predecessor to *Ensueño* (Dream), from the album Barcelona by Montserrat Caballé and Freddie Mercury.

the way we work with syntax and languages, and renewing research into fully automatic parser generation. I wish I discovered earlier that this is what I wanted to attack (and for example had spent less time implementing a Java typechecker), but well, I suppose a thesis project is never finished.

Finally, I would like to emphasize that this area ended up being the subject of my thesis *despite* Eelco Visser's supervision. Since Eelco designed the current revision of SDF and integrated scannerless parsing and generalized LR parsing, this thesis almost looks like a natural continuation of his work. Surprisingly, Eelco continuously encouraged me to move beyond parsing and syntax, but despite his best efforts my work moved closer and closer to parsing issues.

## ACKNOWLEDGEMENTS

*Supervisors*

First of all I thank Eelco Visser, my *'co-promotor'*, for being my supervisor. He was not only my supervisor, but also a mentor and a friend. Almost everything I know about research, I learned from him. Our collaboration was absolutely brilliant. We always worked together on all my publications. The balance between supervising me and contributing to the work we did was perfect. I appreciate how Eelco always gives his students the opportunity to discover things on their own and express their own ideas, as opposed to just telling everything he knows immediately.

I do not remember any formal meeting. We would just end up talking about work every time we wanted to. This must have been difficult sometimes, due to the constant context switches a supervisor has to go through. I am impressed by the capability of some researchers to come up with brilliant remarks just after such a context switch.

Eelco Visser also taught me a lot about publishing strategies. I learned how to target a topic to a general audience, which is most important if you are in a certain niche and your motivation will not be immediately obvious to other researchers. I feel well-equipped to start working as an independent researcher, which is entirely due to Eelco's teaching. Also, Eelco's excellent network gave me the opportunity to meet many interesting researchers, learn how to review papers, and even resulted in finding my future wife. How can one possibly have more impact?

Although I will be moving on to a different country and different universities, I hope we will continue to work together now and then.

I also thank Doaitse Swierstra, my *'promotor'*. Although we didn't meet frequently, I was always struck how well aware he was of our work and my situation. His advice was much appreciated. Also, I have learned that I should probably never buy a cottage for holidays.

*Peers*

I thank the members of the reading committee Mark van den Brand, Dick Grüne, Johan Jeuring, Kees Koster, and Oege de Moor for reviewing my thesis.

Many anonymous reviewers of conferences provided useful suggestions to improve my work. Also, several people voluntarily provided useful feedback on one or more of my papers: Eelco Dolstra, Jeff Gray, Shan Shan Huang, Merijn de Jonge, Karl Trygve Kalleberg, Emmanuel Onzon, and Martijn Vermaat.

All the work of this thesis has been done in collaboration with other people. I thank the co-authors of my publications for their contributions: Eric Bouwers, Arthur van Dam, Eelco Dolstra, René de Groot, Karl Trygve Kalleberg, Koen Muilwijk, Karina Olmos, Éric Tanter, Rob Vermaas, Jurgen Vinju, and Eelco Visser.

*Friends*

Arthur van Dam was always happy to help me with LaTeX and thesis design issues. There must be some contribution from him at almost every single page. He also suggested the Dutch translation *exercities* for *exercises*, which is so much more appropriate than the original *oefeningen*! I will miss speed skating together and even more will I miss our road bicycle racing adventures across the country. How am I supposed to deal with the wind without you? I am sorry I beat you at the only serious hill we climbed together (he still claims he just let me win). I should make sure you get your sweet revenge.

Eelco Dolstra provided me the source of his excellent thesis, which saved me a lot of time during the final preparations of my thesis. He was also a very pleasant officemate. I miss his disquisitions and his exclamations of delight when he has found more remarkable *trivia* at Wikipedia. Eelco also carefully reviewed the Dutch abstract of this thesis.

My friend and ex-colleague Rob Vermaas was always annoying and funny in his own special way. TraCE, a grammar for Shell, and Unicode support are only a few of the things he would happily bother me with almost every week. I will miss the excellent pies prepared occasionally by Lizi Vermaas at the occasion of an accepted paper!

This thesis would not be complete without mentioning the IRC channel whose name must not be mentioned. One of the members, Armijn Hemel, deserves special mentioning for his unique approach towards friendship, but I have no clue what to say about him here. Let me just slap him.

*Logistics*

I pulled quite a few all-nighters (also known as a *bravo*) and I do not think I could have done this without listening to the music of Blank & Jones, the Pet Shop Boys, Moby, and Freddie Mercury. I highly recommend everyone straightforward, gay dance music to get through a long night. I also recommend tea, not coffee. Yunnan green tea works for me.

My loyal computer Logistico could not deal with the finalization of my thesis and the prospect of being abandoned. She died just after submitting my thesis to the reading committee. I owe her every single byte I produced for this thesis.

*Family*

I thank my parents, my sister, and my brother-in-law for their continuous support. Finally, there is Shan Shan, my lovely fiancée. I do not know what to say here about this miracle. Anything I would say about you would be insufficient. I am looking forward to spending my life with you.

Martin Bravenboer
December 6, 2007
Delft