Master Thesis

Syntactic language extensions for ECMAScript

Author:
Matthisk Heimensen
m@tthisk.nl

Supervisor:
dr. Tijs van der Storm
storm@cwi.nl

May 2015

Host organization: Centrum Wiskunde & Informatica http://www.cwi.nl

Universiteit van Amsterdam
Faculteit der Natuurwetenschappen, Wiskunde en Informatica
Master Software Engineering
http://www.software-engineering-amsterdam.nl

UNIVERSITEIT VAN AMSTERDAM

Abstract

Faculteit der Natuurwetenschappen, Wiskunde en Informatica
Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Master of Science

Syntactic language extensions for ECMAScript

by Matthisk HEIMENSEN

The Thesis Abstract is written here (and usually kept to just this page). The page is kept centered vertically so can expand into the blank space above the title too...

Acknowledgements

The acknowledgements and the people to thank go here, don't forget to include your project advisor…

Contents

Abbreviations

ES6	**E**CMA**S**cript 6
ES5	**E**CMA**S**cript 5
LOC	**L**ines **O**f **C**ode
AST	**A**bstract **S**yntax **T**ree
HOAS	**H**igher **O**rder **A**bstract **S**yntax

Chapter 1

Introduction and Motivation

1.1 Context

JavaScript is an implementation of the ECMAScript specification. Since the release of the first version of the specification there have been four iterations (where version four was skipped), the current version of the specification is ECMAScript 5. In the near future a new version of the specification will leave the draft status and become the standard (ECMAScript 6). With a new specification come new syntactic language features. Before these features can be used in the JavaScript run-time environment of choice, the vendors of these environments have to implement the new standard. However it is possible to use the new standard today through the use of program transformations.

In this work we study the extension of programming languages, the new ECMAScript specification presents an opportunity to research these extensions and there transformations. Using the RASCAL meta-programming environment[1] we will implement the program transformations. Afterwards we can perform an analysis on the resulting transformation suite to uncover the *effectiveness* of our implementation.

"Program transformations find ubiquitous application in compiler construction to realize desugaring, optimizes, and code generators"[2] in our research we will focus on the desugaring of (new) language constructs.

"The aim of program transformation is to increase programmer productivity by automating programming tasks, thus enabling programming at a higher-level of abstraction, and increasing maintainability"[3]

1.2 Problem definition

Program transformations are faced by multiple problems, in what order do we run the transformation rules, how do we represent the program (e.g. abstract syntax) to perform transformations on, what guarantee do we have for the validity of our transformations, how can a transformation introduce new bindings. Some of these problems reoccur in every transformation that is created. These problems are identified as cross-cutting concerns. The goal of this thesis is to identify these concerns and separate them from the transformation code.

The capture avoidance is one such problem, it is mostly studied in the context of macro expansion[4–6]. And is often called the hygiene of transformations. When an identifier originating from the source program references to a synthesized declaration, the transformation is sub-hygienic and introduces variable capture. To prevent this problem we implement the *name-fix* algorithm[2] for our transformation suite.

1.3 Scope

1.3.1 Expected results

This thesis will have the following results:

- A taxonomy of ES6 language features specified by their needed transformations
- Language extension suite implementing (a subset of) the ES6 features.
- A set of compatibility tests, testing our implementation against the specification document.
- Eclipse integration (syntax highlighting & declared at hyperlinks) for new language features

1.4 Research questions

The central research questions of this thesis are:

1. How can independent program transformations be categorized?
2. What are the cross-cutting concerns of any program transformation, and how can we take care of these concerns outside of the transformations themselves?
3. What is the advantage of using the RASCAL language workbench[1] for the creation of our transformation suite?

1.5 Outline

In Chapter 2 the background (program transformations, language workbench) is introduced. In Chapter 3 we give an overview of language extensions that are introduced and make a taxonomy of their transformations. In **??** we discuss the cross-cutting concerns of the transformation suite (i.e. variable capture). In Chapter **??** we evaluate our resulting transformation suite (against others). Finally, we conclude in Chapter **??**.

Chapter 2

Background

2.1 Program Transformations

Eelco Visser defines the aim of a program transformation as follows:

"The aim of program transformation is to increase programmer productivity by automating programming tasks, thus enabling programming at a higher-level of abstraction, and increasing maintainability" ([3])

There are many different types of program transformations. Programs can be transformed from a source to a target language. Where both can be the same or different. One can be a higher level language and one a lower level language. In this thesis we will focus on transformations in the category rephrasing[3]. Here a program in a source language is transformed to a different program in the same language. The use case in our thesis is language extensions see section 2.2.

2.2 Language Extensions

"Language extensions augment a base language with additional language features. Many compilers first desugar a source program to a core language."[2] For example the Haskell functional programming language defines for many constructs how they can be transformed to a kernel language[7]

2.3 Program Representation

?? Before programs can be transformed they need a structured representation. Seldom are program transformations performed on the merely the textual input. Here we discuss some program representations.

Parse Trees Parse trees represent a programs syntactic information in a tree. This tree includes layout information of the input program (e.g. white-space, comments). The parse tree is structured according to some context-free grammar, that defines how to parse textual input. The parse tree also deals with disambiguation of input and representation of parentheses.

Abstract Syntax Trees or AST is created when layout information is removed from a parse tree and only the core constructs of the language grammar remain as nodes in the tree. White-space layout, comments, and things like parentheses (i.e. implicitly implied by the structure of the AST) are removed. An AST is often the input for compiler pipelines, program transformations, and refactorings.

Higher Order Syntax Trees (HOAS) To represent not just a program's sub expression relations but also its variable binding we can use a Higher Order Syntax Tree (HOAS)[**?**]. In a HOAS the variable bindings are made explicit (just as the sub expression relations are made explicit in an AST). Every variable has a binding-site and possibly uses throughout the rest of the tree. *"In addition to dealing with the problem of variable capture, HOAS provides higher-order matching which synthesizes new function for higher-order variables. One of the problems of higher-order mathcing is that there can be many matches for a pattern"*[3]

2.4 The Language Workbench

The language workbench is a term popularized by Martin Fowler and we can formulate his definition as follows:

A language workbench makes it easy to build tools that match the best of modern IDEs, where the primary source of information is a persistent abstract representation. Its users can freely design new languages without a semantic barrier. With the result of language oriented programming becoming much more accessible.[8]

> Essentially the promise of language workbenches is that they provide the flexibility of external DSLs without a semantic barrier. Furthermore they make it easy to build tools that match the best of modern IDEs. The result makes language oriented programming much easier to build and support, lowering the barriers that have made language oriented programming so awkward for so many. ([8])

These workbenches reduce the awkwardness of language oriented programming[9] significantly, by giving programmers the the tools to define programming languages with an abstract representation.

The RASCAL[1] metaprogramming environment (mpl) is a language workbench. It allows programmers to define context-free grammars, generate parsers for these grammars. With these parsers programmers are able to define parse tree patterns. The RASCAL mpl uses Eclipse for IDE integration, which allows programmers to define interaction handles and syntax highlighting. All language extensions presented in this thesis are implemented using the RASCAL mpl.

Concrete syntax A Rascal generated parser returns a parse tree, which is an ordered, rooted tree that represents the syntactic structure of a string according to the formal grammar used to specify the parser. Most transformation systems would implode this concrete syntax tree to an AST and perform program transformations on this tree. Rascal gives the possibility to perform program transformation directly on the concrete syntax tree through the use of concrete-syntax patterns. Presenting multiple advantages over an AST based solution:

- Preserving layout information encoded in the concrete-syntax tree
- Avoiding the use of a pretty printer to output the transformed AST to a textual representation
- Transformation code reads as rewrite rules, instead of a clutter of AST nodes.

A concrete-syntax pattern may contain variables in the form of our grammars non-terminals, these variables are bound in the context of the current patterns use (e.g. functions body when used as a formal parameter). The following pattern matches un-named JavaScript functions according to the Saner JavaScript syntax definition[1]

```
(Expression)`function( <Params ps> ) { <Statement* body> }` := pt
```

In similar fashion new syntax trees can be constructed, using the variables bound from pattern matches.

```
(Expression)`function plus(x, y) { return x + y; }`
```

Rascal also supplies a shorthand notation to parse a string starting from the supplied non-terminal

```
[Expression]"1 + 2"
```

[1]https://goo.gl/B6HR22

Chapter 3

Transforming ECMAScript

3.1 Motivating Examples

The Babeljs compiler visits each AST node, functions inside of a specific transformation can announce themselves for the callback stack of a node visit. The callback to this function receives several arguments (current AST node, parent AST node, current Scope, and the file being transformed). The Scope class has functions presented in 3.1, if a node transformation wants to introduce a new variable say ref it calls $generateUidIdentifier$ with String ref as an argument on Scope and receives a name which is not bound in the current Scope. For this to work the visitor needs to have information of all the variables bound in the current scope. This makes the transformation more a full compiler than a set of transformation rules. For starters the transformation code needs to be aware of the Scope class and its functions.

```
generateUidIdentifier(name: string) {
  return t.identifier(this.generateUid(name));
}

generateUid(name: string) {
  name = t.toIdentifier(name).replace(/^_+/, "");

  var uid;
  var i = 0;
  do {
    uid = this._generateUid(name, i);
    i++;
  } while (this.hasBinding(uid) || this.hasGlobal(uid) || this.hasReference(uid));


  var program = this.getProgramParent();
  program.references[uid] = true;
  program.uids[uid] = true;

  return uid;
}

_generateUid(name, i) {
  var id = name;
  if (i > 1) id += i;
  return `_${id}`;
```

```
  }
```

Listing 3.1: Variable capture avoidance code from babeljs source[1]

The Traceur compiler totally ignores the problem of variable capture. In conformance with the case study conducted by Erdeweg et. al. [2] we were able to identify a transpiler that fails to address variable capture. With a simple example we can demonstrate the introduction of capture by Traceur 3.2

```
function f() {
  var $__0 = 0;
  var x = () => this;
}
```

Listing 3.2: Example input to Traceur[2]

```
function f() {
    var $__0 = this;
    var $__0 = 0;
    var x = (function() {
      return $__0;
    });
  }
```

Listing 3.3: Variable capture

3.2 Taxonomy of language features

Every language extension has several properties which can be identified and categorized along certain dimensions. In this chapter we present such a taxonomy for all the language extensions presented in the ECMAScript 6 specification[10].

3.2.1 Dimensions

The following dimensions are identified and used to categorize every language extension.

Category One of the rephrasing categories defined by Eelco Visser[3]:

- Simplification
- Desugaring
- Weaving
- Specialization
- Inlining

[1] https://goo.gl/BZKIvV
[2] https://goo.gl/Ds3xUn

- Fusion
- Design improvement
- Obfuscation
- Renovation

Abstraction level Program transformations can be categorized by their abstraction level. There are four levels of abstraction (similar to those of macro expansions[**?**]), character-, token-, syntax-, or semantic-based. Character and token based transformations work on a program in textual representation. Syntactical transformations work on a program in its parsed representation (either as an AST or as a parse tree **??**). Next to the syntactic representation semantic transformations also have access to the static semantics of the input program (e.g. variable binding).

Purely semantic Is the extension purely in semantic, or does the extension introduce new syntax.

Scope Program transformations can happen within four different scopes:

When a program transformation matches on a sub-tree of the parse tree and only transforms this matched sub-tree it is a local-to-local transformation. If the transformation needs information outside the context of the matched sub-tree, but only transforms the matched sub-tree it is global-to-local. When a transformation has no additional context from its local sub-tree but does alter the entire parse tree it is called local-to-global.If the transformation transforms the input program in its entirety it is global-to-global.

Syntactically type preserving Program transformations performed on syntax elements (i.e. being syntax- or semantic based) can preserve the syntax type of their input element or alter it (e.g. is an expression transformed to an expression, or to a list of statements)

Introduction of bindings Are new bindings introduced in the transformed code as opposed to the original code.

Depending on bindings (i.e. run-time code) Will the transformed code rely on function calls not introduced by the transformation itself but provided separately from the transformation suite.

Compositional ...

Preconditions What are the preconditions that have to be met before execution of a transformation rule, to ensure validness of our transformation (e.g. all sub-terms have to be analyzed and transformed)

Restrictions on sub-terms Does the language extensions impose restrictions on the sub-terms used inside of the language extension.

Analysis of sub-terms Are the non-terminals of our language extension analyzed and possibly transformed by the transformation rule.

Dependency on other extensions Can the language extensions be performed standalone or is there a dependency on one of the other extensions.

Backwards compatible Is the API of the transformed code compatible with the ECMAScript 6 specification (i.e. can we import a transformed module in ECMAScript 6 and use it properly).

Dividable Is it possible to identify smaller transformation rules inside this language extension, that can be performed independently from one another.

3.2.2 Taxonomy

3.2.2.1 Arrow Functions

Arrow functions[10, 14.2] are the new lambda-like function definitions inspired by Coffeescript and C# notation. The functions body knows no lexical this binding but instead uses the binding of its parent lexical scope.

Table 3.1: Extension transformation dimensions

	Arrow Functions
Category	Desugaring
Transformation level	Context-free syntax
Scope	Global-to-local
Syntactically type preserving	Yes
Introducing bindings	Yes
Depending on bindings	No
Compositional	Yes
Analysis of sub-terms	Yes
Constraints on sub-terms	Yes
Preconditions	Yes
Dependencies	No
Backwards compatible	Yes
Dividable	No

Category The arrow function construct is eliminated by translating them into the fundamental function construct. Thus we speak of a syntactic sugar and a desugaring.

Scope To avoid a ReferenceError when an arrow function is used outside of the lexical scope of a function, the $_{a}rguments$ identifier should reference $undefined$. But only when the arrow is placed outside of any functions lexical scope. Thus the transformation needs context information of the placement of an arrow function and the transformations scope is global-to-local.

Analysis of sub-terms References to the arguments, sup er and this keyword have to be identified in the ConciseBody of an ArrowFunction and renamed.

"An ArrowFunction does not define local bindings for **arguments**, **super**, **this**, or new.target. Any reference to arguments, super, or this within an ArrowFunction must resolve to a binding in a lexically enclosing environment." ([10, 14.2.16])

An arrow function is transformed to an ES5 function which is wrapped in a self-calling function which receives its lexical enclosing scopes **super**, **this**, and **arguments** variables. The enclosing lexical scope is not polluted with new variable declarations. Keywords references are prepended with an underscore, only references in the current scope (i.e. no deeper scope) are renamed.

```
var f = <Arrow Function>;
```

```
var f = (function(_super,_arguments,_this) {
  return <Desugared Arrow Function>
})(super,arguments,this);
```

Constraints on sub-terms Use of the **yield** keyword is not allowed in an arrow functions ConciseBody. Arrow functions can not be generators and deep continuations are avoided.

Preconditions Before an arrow function can be transformed its ConciseBody should have been transformed. This to prevent the incorrect renaming of sub-terms (**super**,**arguments**,**this**), because they still fall in the same scope depth.

```
var f = () => {
  () => this;
};
```

```
var f = (function(_this,_arguments) {
  return function() {
    () => _this;
  };
})(this,arguments);
```

Listing 3.4: Incorrect renaming of this in nested arrow function

3.2.2.2 Classes

Class definitions[10, 14.5] are introduced in ECMAScript 6 as a new feature to standardize inheritance model. Underneath the prototypal inheritance model is still used to create class declarations.

Table 3.2: Extension transformation dimensions

	Classes
Category	Desugaring
Abstraction level	Context-free syntax
Scope	Local-to-local
Syntactically type preserving	Yes
Introducing bindings	No
Depending on bindings	Yes
Compositional	Yes
Analysis of sub-terms	Yes
Constraints on sub-terms	Yes
Preconditions	Yes
Dependencies	No
Backwards compatible	Yes
Dividable	No

Category The class construct is syntactic sugar for simulation of class based inheritance through the more fundamental prototypal inheritance system of the ECMAScript language.

Syntactically type preserving There are two types of class declarations, a direct declaration as a Statement, or an Expression declaration. They are both transformed to a construct of the same syntactical type.

Depending on bindings Several features of the class declaration as described in the specification demand some run-time. (e.g. the class methods being non enumerable)

Analysis of sub-terms References of **super** and a **super** call have to identified inside constructor or class methods. These are transformed to preserve the correct **this** binding.

Constraints on sub-terms The sub-terms of a class declaration are always executed in strict mode

“A ClassBody is always strict code.” ([10, 14.5])

Preconditions Bodies of constructor and class methods should have been transformed before the class declaration itself is transformed. This to prevent incorrect transformation of the sub-term **super**.

3.2.2.3 Destructuring

Destructuring[10, 12.14.5] is a new language construct to extract values from object or arrays with a single assignment. It can be used in multiple places among which parameters, variable declaration, and expression assignment.

TABLE 3.3: Extension transformation dimensions

	Destructuring
Category	Desugaring
Abstraction level	Context-free syntax
Scope	Local-to-local
Syntactically type preserving	No
Introducing bindings	Yes
Depending on bindings	Yes
Compositional	Yes
Analysis of sub-terms	Yes
Constraints on sub-terms	No
Preconditions	No
Dependencies	Yes
Backwards compatible	Yes
Dividable	Yes

Category The destructuring language feature is eliminated by translating it into fundamental language concepts such as member access on objects and array member selection.

Syntactically type preserving The transformation is not syntactically type preserving in every situation. If the destructuring is used in a variable declaration the syntactic type is converted from $< Statement >$ to $< Statement* >$, because new bindings are introduced.

Dependencies Object destructuring supports the computed property notation of extended object literals (as discussed in section: 3.2.2.4):

```
var qux = "key";
var { [qux] : a } = obj;
```

Destructuring thus has a dependency on extended object literal notation for this feature to work properly.

Dividable This language features consists of two types of destructuring, object destructurings, and array destructurings. These two different features can be transformed separately.

3.2.2.4 Extended object literals

Literal object notation receives three new features in the ECMAScript 6 standard[10, 12.2.5]. Shorthand property notation, shorthand method notation, and computed property names.

TABLE 3.4: Extension transformation dimensions

	Object literals
Category	Desugaring
Abstraction level	Context-free syntax
Scope	Local-to-local
Syntactically type preserving	Yes
Introducing bindings	No
Depending on bindings	No
Compositional	Yes
Analysis of sub-terms	No
Constraints on sub-terms	No
Preconditions	No
Dependencies	No
Backwards compatible	Yes
Dividable	Yes

Category The extended object literals are syntactic sugar in the purest form. There are direct rules to eliminate the introduced concepts to fundamental concepts of the ECMAScript 5 language. (e.g. shorthand property notation translate to non-shorthand notation)

Introducing bindings For full spec compliance run-time code has to be introduced to set computed property keys. However this can also be done through the member notation:

```
var obj = { [qux] : 0; }
```

```
var _obj;
var obj = ( _obj = {}, _obj[qux] = 0, _obj );
```

Dividable All three extensions of the object literal can be transformed separately.

3.2.2.5 For of loop

The ECMAScript 6 standard introduces iterators, and a shorthand for notation to loop over these iterators in the form of for of[10, 13.6.4]. Previous versions of the ECMAScript standard had default for loops, and for in loops (which iterate over all enumerable properties of an object).

TABLE 3.5: Extension transformation dimensions

	For of loop
Category	Desugaring
Abstraction level	Context-free syntax
Scope	Local-to-local
Syntactically type preserving	Yes
Introducing bindings	No
Depending on bindings	No
Compositional	Yes
Analysis of sub-terms	No
Constraints on sub-terms	No
Preconditions	No
Dependencies	Yes
Backwards compatible	Yes
Dividable	Yes

Category This construct can be eliminated by transformation to a for-loop using an iterator variable, and selecting the bound variable for the current index from the array using this iterator variable.

Constraints on sub-terms No constraints any $< \mathit{Statement} >$ can be used within the body of a for-of loop.

```
syntax Statement
  = "for" "(" Declarator ForBinding "of" Expression ")" Statement;
```

Dependencies The for-of loop binding can be declared with the let declarator**??**.

```
for( let x of arr ) <Statement>
```

3.2.2.6 Spread operator

ECMAScript 6 introduces a new unary operator named spread[10, 12.3.6.1]. This operator is used to expand an expression in places where multiple arguments (i.e. function calls) or multiple elements (i.e. array literals) are expected.

Category This unary operator can be expressed using fundamental concepts of the ECMAScript 5 language. In case the operator appears in a place where multiple arguments are expected, a call of member function *apply* on our function can be used to supply arguments as an array. In case it appears in a place where multiple elements are expected the prototype function *concat* of arrays can be used to interleave the supplied argument to the operator with the rest of the array.

Table 3.6: Extension transformation dimensions

	Spread operator
Category	Desugaring
Abstraction level	Context-free syntax
Scope	Local-to-local
Syntactically type preserving	Yes
Introducing bindings	Yes
Depending on bindings	Yes
Compositional	Yes
Analysis of sub-terms	Yes
Constraints on sub-terms	No
Preconditions	No
Dependencies	No
Backwards compatible	Yes
Dividable	Yes

Introducing bindings A binding has to be introduced to save the scope in which a function is applied, when the function call happens on an object3.2.2.6

```
a.b.f(...args);
```

```
var _a$b;
(_a$b = a.b).f.apply(_a$b, args);
```

Depending on bindings For correct specification compliance the argument of the spread operator has to be transformed to a consumable array. This behavior can be simulated through a run-time function.

Analysis of sub-terms If there exists a function call of function f on object a where spread operator is used in arguments list. This function call has to be transformed to a call to function *apply* on function f on object a. Where the first parameter of *apply* is a and second is the argument of spread. The sub-terms of function calls on objects with spread operators thus have to be analyzed to determine the call scope of the called function.

```
a.f(...args);
```

```
a.f.apply(a,args);
```

Listing 3.5: apply function with correct *this* scope

Dividable The operator can be used in two places (places of multiple arguments, and multiple elements), which can be transformed separately.

3.2.2.7 Default parameter

The ECMAScript 6 spec defines a way to give parameters default values[10, 9.2.12]. These values are used if the caller does not supply any value on the position of this argument. Any default value is evaluated in the scope of the function (i.e. *this* will resolve to the run-time context of the function the default parameter value is defined on).

Table 3.7: Extension transformation dimensions

	Default parameters
Category	Desugaring
Abstraction level	Context-free syntax
Scope	Local-to-local
Syntactically type preserving	Yes
Introducing bindings	No
Depending on bindings	No
Compositional	Yes
Analysis of sub-terms	No
Constraints on sub-terms	No
Preconditions	No
Dependencies	No
Backwards compatible	Yes
Dividable	No

Category This language feature can be eliminated by transformation to a fundamental concept of the ECMAScript 5 language. By binding the variable in the function body to either the default value (in case the parameter equals $undefined$) or to the value supplied by the parameter.

Introducing bindings This transformation does not introduce any bindings because the bindings already exist as formal parameters.

3.2.2.8 Rest parameter

The *BindingRestElement* defines a special parameter which binds to the remainder of the arguments supplied by the caller of the function.[10, 14.1]

Category

Syntactically type preserving $< Function >$ to $< Function >$

Table 3.8: Extension transformation dimensions

	Rest Parameter
Category	Desugaring
Abstraction level	Context-free syntax
Scope	Local-to-local
Syntactically type preserving	Yes
Introducing bindings	No
Depending on bindings	No
Compositional	Yes
Analysis of sub-terms	No
Constraints on sub-terms	No
Preconditions	No
Dependencies	No
Backwards compatible	Yes
Dividable	No

3.2.2.9 Template strings

Standard string literals in JavaScript have some limitations. The ECMAScript 6 specification introduces a template string literal to overcome some of these[10, 12.2.8]. Template string literals are delimited by the ‘ quotation mark, can span multiple lines and be interpolated by expressions: $\$\{< Expression >\}$

Table 3.9: Extension transformation dimensions

	Template Strings
Category	Desugaring
Abstraction level	Context-free syntax
Scope	Local-to-local
Syntactically type preserving	Yes
Introducing bindings	No
Depending on bindings	No
Compositional	Yes
Analysis of sub-terms	No
Constraints on sub-terms	No
Preconditions	No
Dependencies	No
Backwards compatible	Yes
Dividable	No

Category Template strings are syntactic sugar for Strings concatenated with expressions and new-line characters.

Syntactically type preserving $< TemplateLiteral >$ to $< StringLiteral >$

3.2.2.10 Tail call optimization

JavaScript knows no optimization for recursive calls in the tail of a function, this results in the stack increasing with each new recursive call overflowing the stack as a result. The ECMAScript 6 specification introduces tail call optimization to overcome this problem[10, 14.6]

Category Optimization, this transformation improves the space performance of a program.

Abstraction level Tail call optimization is a semantic based transformation. To identify if the expression of a return statement is a recursive call to the current function we need name binding information.

Pure semantics This extension is purely semantic, no new syntax is introduced. To transform this extension to work in older JavaScript interpreters the tail call has to be removed and replaced by a *while* loop, to prevent the call stack from overflowing.

```
function fib(n,nn,res) {
  if( nn > 100 ) return res;
  return fib(nn,n + nn,res + [n,nn]);
}
```

Listing 3.6: Function with tail recursion

```
function fib(arg0,arg1,arg2) {
  while (true) {
    var n = arg0, nn = arg1, res = arg2;

    if(nn > 100) return res;
    arg0 = nn;
    arg1 = n + nn;
    arg2 = res + [n,nn];
  }
}
```

Listing 3.7: Semantically identical function, without tail recursion

Figure 3.7 illustrates how a function with tail recursion can be transformed to a function with an infinite loop. Not filling the call stack with new return positions and thus executing correctly even when our upper limit (100) is increased.

3.2.2.11 Generators

The ECMAScript 6 standard introduces a new function type, called generators[10, 14.4]. These are functions that can be exited and later re-entered, when leaving the function through *yielding* their context (i.e. variable bindings) is saved and restored upon re-entering the function. A generator function is declared through `function*` and returns a Generator object (which inherits from *Iterator*)

Table 3.10: Extension transformation dimensions

	Generators
Category	Desugaring
Abstraction level	Context-free syntax
Scope	Local-to-local
Syntactically type preserving	Yes
Introducing bindings	No
Depending on bindings	Yes
Compositional	Yes
Analysis of sub-terms	Yes
Constraints on sub-terms	No
Preconditions	Yes
Dependencies	No
Backwards compatible	Yes
Dividable	No

Category ...

Scope Including a run-time the transformation can be managed local-to-local.

Syntactically type preserving $< \mathit{Statement} >$ to $< \mathit{Statement} >$ $< \mathit{Expression} >$ to $< \mathit{Expression} >$

Analysis of sub-terms Identifying the use of *yield* and *return*

Preconditions The $< \mathit{GeneratorBody} >$ has to be transformed before transformation of the Generator function starts.

Depending on bindings A run-time system can be used to simulate the behavior of Generator functions[3]

3.2.2.12 Let and Const declarators

JavaScript's scoping happens according to the lexical scope of functions. Variable declarations are accessible in the entire scope of the executing function, this means no block scoping. The *let* and *const* declarator of the ECMAScript 6 specification change this. Variables declared with either one of these is lexically scoped to its block and not its executing function. The variables are also not hoisted to the top of their block, they can only be used after they are declared.

[3] http://github.com/facebook/regenerator/blob/master/runtime.js

TABLE 3.11: Extension transformation dimensions

	Let Const
Category	Undefined
Abstraction level	Semantic
Scope	Global-to-global
Syntactically type preserving	Yes
Introducing bindings	No
Depending on bindings	No
Compositional	No
Analysis of sub-terms	Yes
Constraints on sub-terms	No
Preconditions	Yes
Dependencies	No
Backwards compatible	No
Dividable	Yes

Category This language feature does not fall under any of the defined categories of rephrasing by Eelco Visser[3]. The feature introduces a new language concept, however this concept is not syntactic sugar for a more fundamental concept in ECMAScript 5. This because the previous standard does not know *block-scoping*.

Scope The transformation needs to analyse the entire program (global) and make transformations to its entirety (global). Thus the scope of the transformation is global-to-global.

Compositional The transformation of *let* (or *const*) bindings is not compositional because bindings of these types can not be captured by names in sibling scopes

```
{
  let x = 0;
}
{
  let x = 1;
}
```

Once these *let* bindings are transformed to *var* bindings without renaming they would be declared in the same lexical scope. Whereas they are not at the moment. For a transformation to *var* bindings to work one of these bindings and its uses has to be renamed.

Introducing bindings The transformation introduces no new bindings, it converts *let* and *const* bindings to *var* bindings, and renames these bindings were necessary.

Preconditions Other transformations need to be performed before the let const extension is transformed. To prevent this extension to have dependencies on other extensions. For example the for-of extension has a dependency on the let-const extension:

```
for( let x of arr ) <Statement>
```

If the for-of loop is transformed to a normal loop let-const needs not to be aware of the for-of loop as an extension, but just of the ECMAScript 5 manner to declare variables inside loops.

Dividable Variables declared through let and const can be transformed to ECMAScript 5 code separately.

3.2.3 On the expressive power of ECMAScript 6

3.3 Implementation

Each transformation is defined as a term-rewriting rule. It has a concrete syntax pattern which can match part of a parse-tree. The result is a concrete piece of syntax, using only constructs from the fundamental syntax definition (i.e. ECMAScript 5). The rewrite rules are exhaustively applied on the input parse-tree until no more rewrite rules match any sub-trees of the input. Application of rewrite rules to the parse tree is done bottom-up, because several rewrite rules (e.g. 3.1) demand for successful completion that their sub-terms are already transformed.

Appendix A

Appendix Title Here

Write your Appendix content here.

Bibliography

[1] Paul Klint, Tijs van der Storm, and Jurgen Vinju. RASCAL : a Domain Specific Language for Source Code Analysis and Manipulation.

[2] Sebastian Erdweg, Tijs van der Storm, and Yi Dai. Capture-Avoiding and Hygienic Program Transformations (incl. Proofs). 8586:1–29, April 2014. URL http://link.springer.com/10.1007/978-3-662-44202-9http://arxiv.org/abs/1404.5770.

[3] Eelco Visser. A Survey of Rewriting Strategies in Program Transformation Systems. *Electronic Notes in Theoretical Computer Science*, 57:109–143, December 2001. ISSN 15710661. doi: 10.1016/S1571-0661(04)00270-1. URL http://linkinghub.elsevier.com/retrieve/pii/S1571066104002701.

[4] David Herman and Mitchell Wand. A Theory of Typed Hygienic Macros. *Proceedings of the 17th European Symposium on Programming*, 4960:48, 2010. URL http://www.springerlink.com/index/V03K0734Q7217R35.pdf.

[5] David Herman and Mitchell Wand. A theory of hygienic macros. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 4960 LNCS, pages 48–62, 2008. ISBN 3540787380. doi: 10.1007/978-3-540-78739-6_4. URL http://www.springerlink.com/index/V03K0734Q7217R35.pdf.

[6] Tim Disney, Nathan Faubion, David Herman, and Cormac Flanagan. Sweeten your JavaScript. In *Proceedings of the 10th ACM Symposium on Dynamic languages - DLS '14*, pages 35–44, New York, New York, USA, 2014. ACM Press. ISBN 9781450332118. doi: 10.1145/2661088.2661097. URL http://dl.acm.org/citation.cfm?doid=2661088.2661097.

[7] SL Peyton Jones, RJM Hughes, L Augustsson, D Barton, B Boutel, W Burton, J Fasel, K Hammond, R Hinze, P Hudak, T Johnsson, MP Jones, J Launchbury, E Meijer, J Peterson, A Reid, C Runciman, and PL Wadler. Report on the programming language haskell 98, February 1999. URL http://research.microsoft.com/apps/pubs/default.aspx?id=67041.

[8] Martin Fowler. Language workbenches: The killer-app for domain specific languages. *Accessed online from: http://www. martinfowler. com/articles/languageWorkbench. html*, pages 1–27, 2005. doi: http://www.martinfowler.com/articles/languageWorkbench.html. URL http://www.issi.uned.es/doctorado/generative/Bibliografia/Fowler.pdf.

[9] M P Ward. Language Oriented Programming. *Software-Concepts and Tools*, 15(4): 1–22, 1994.

[10] Ecmascript 2015 language specification. final draft 38, Ecma International, apr 2015.