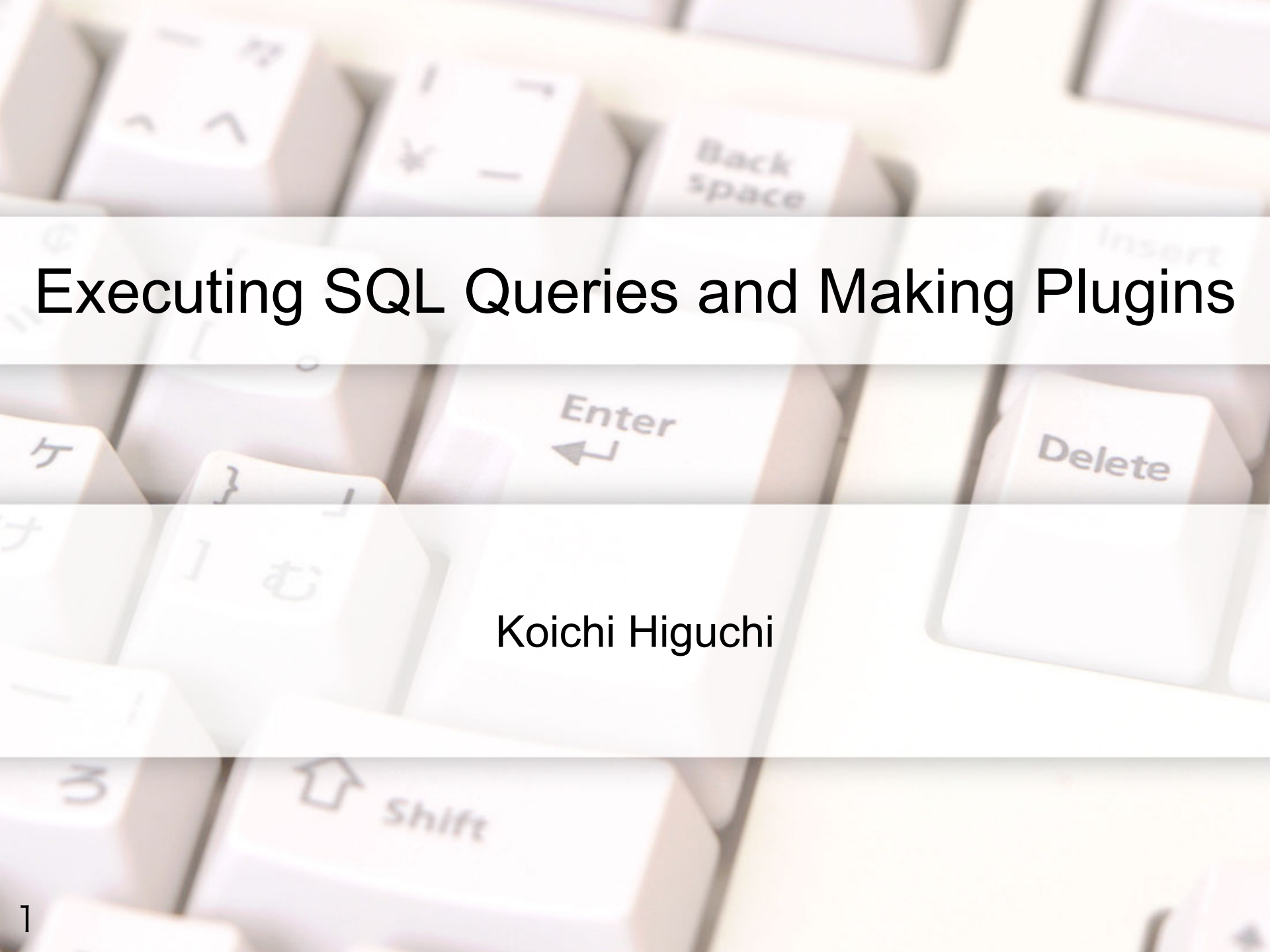
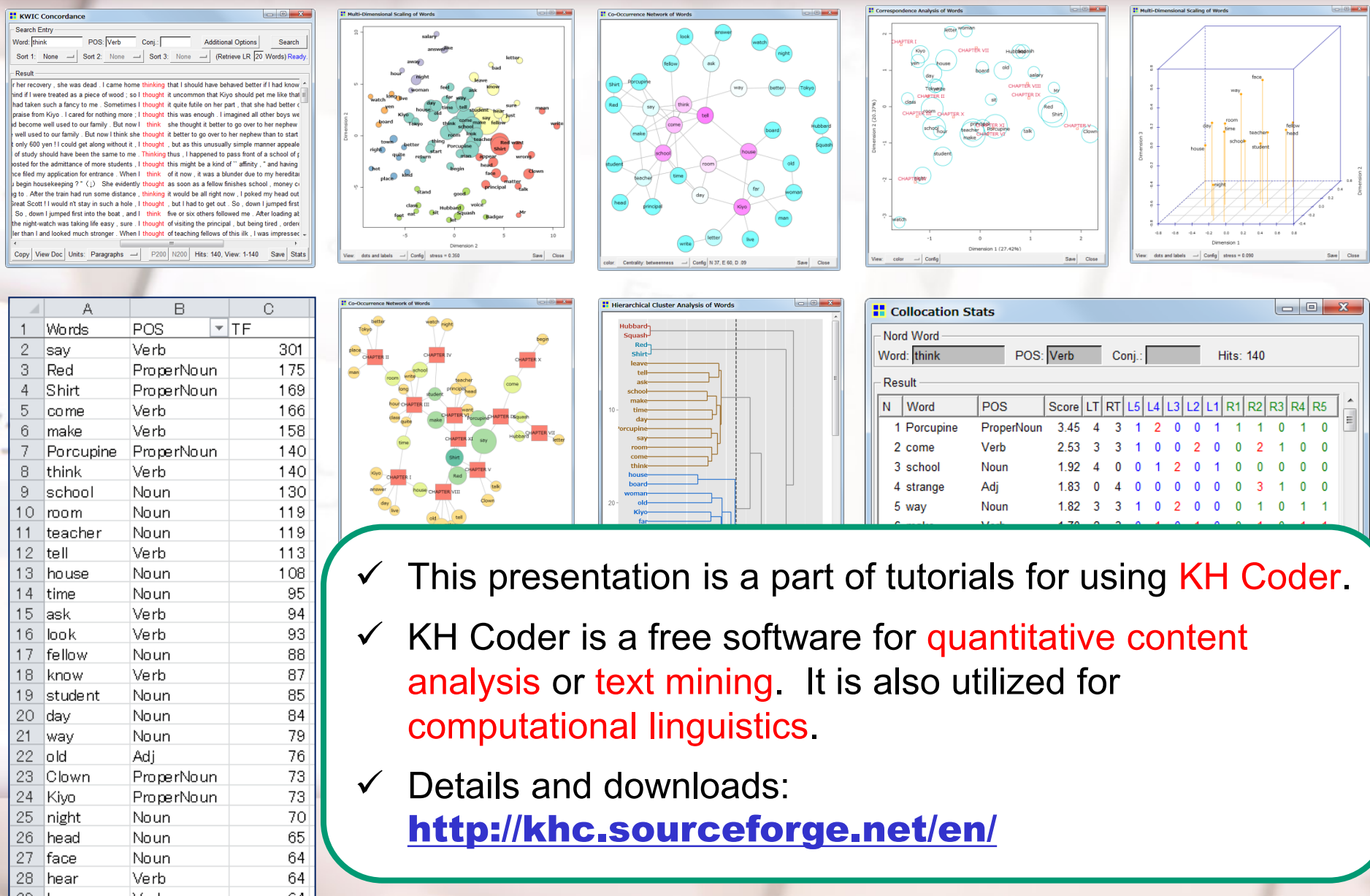




Executing SQL Queries and Making Plugins

Koichi Higuchi

Preface



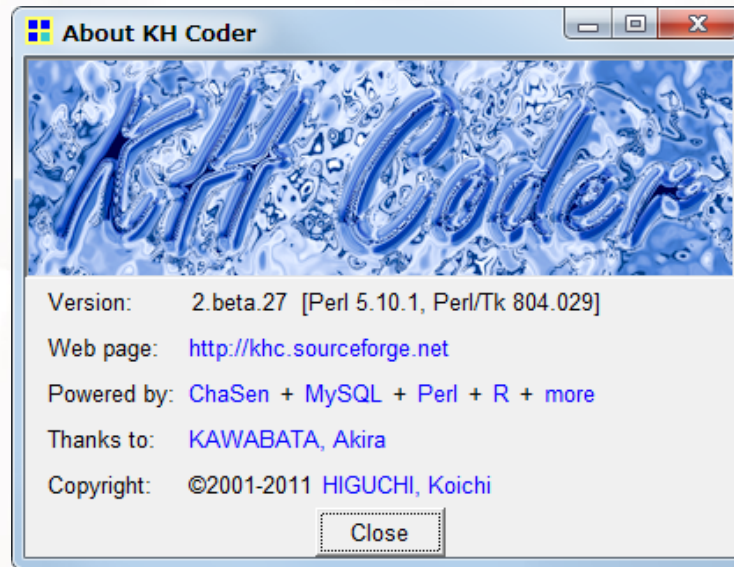
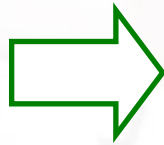
- ✓ This presentation is a part of tutorials for using **KH Coder**.
- ✓ KH Coder is a free software for **quantitative content analysis** or **text mining**. It is also utilized for **computational linguistics**.
- ✓ Details and downloads: <http://khc.sourceforge.net/en/>

First Impression Can Be:

KH Coder



Text

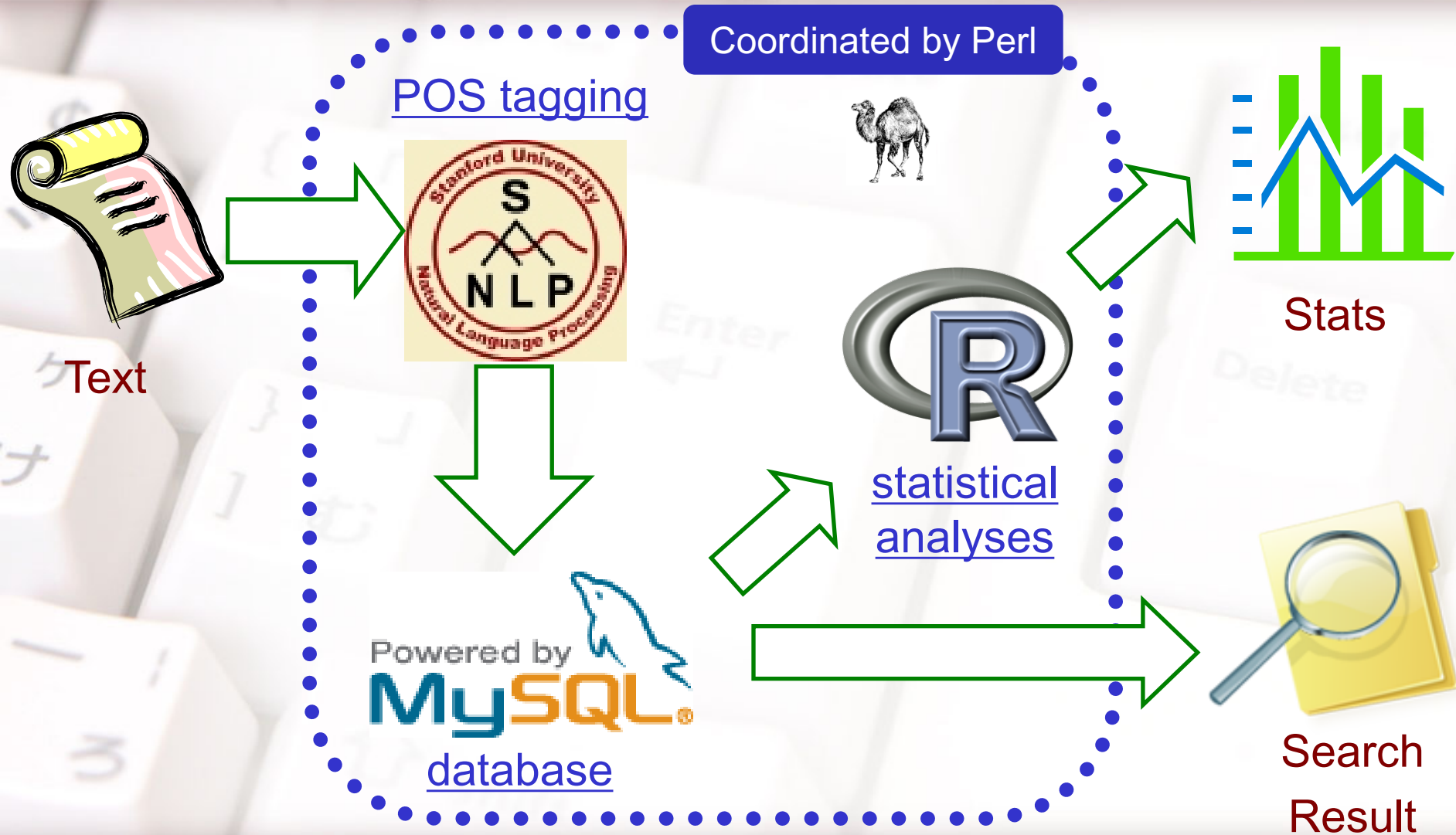


Stats



Search
Result

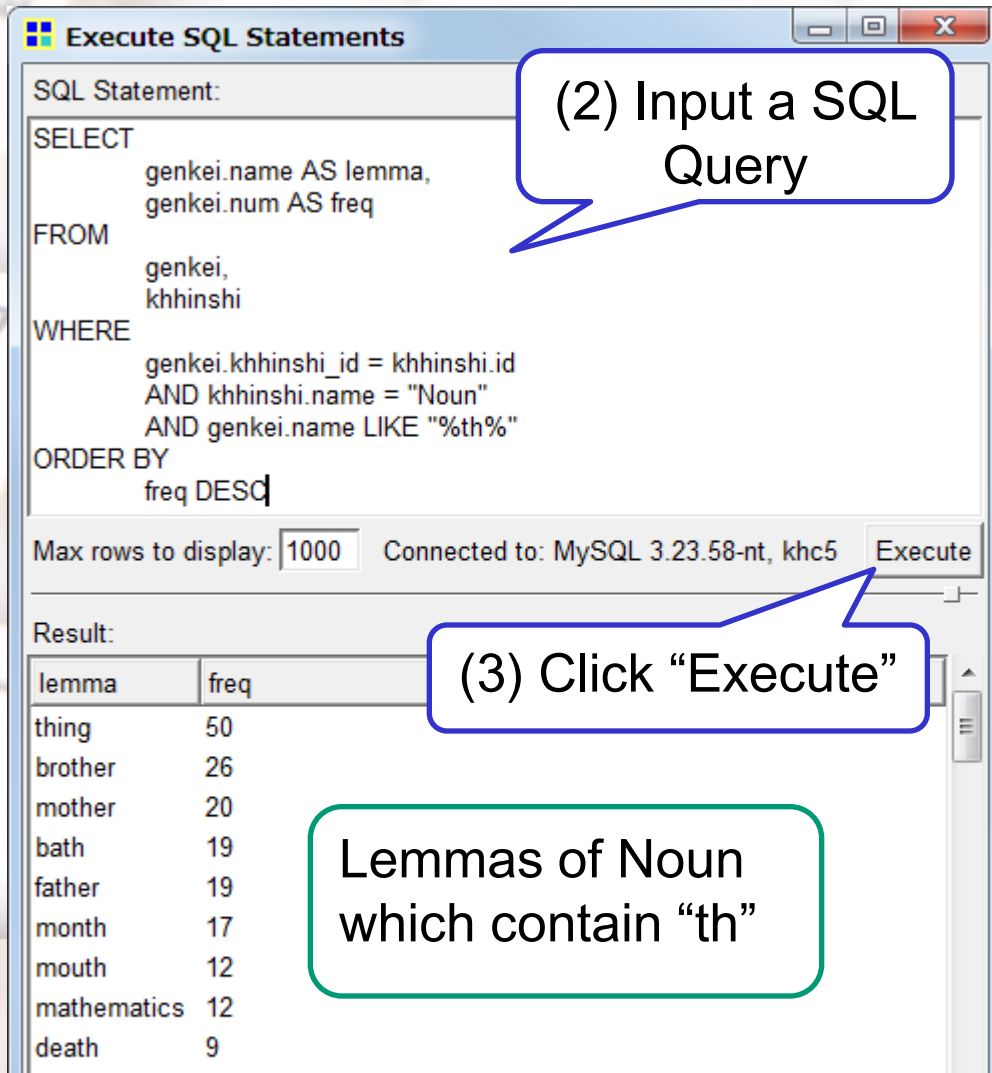
To Be More Specific:



Copyrights of logos on this page belong to their respective owners.
Please click the logos to visit their web sites.

Execute SQL Queries

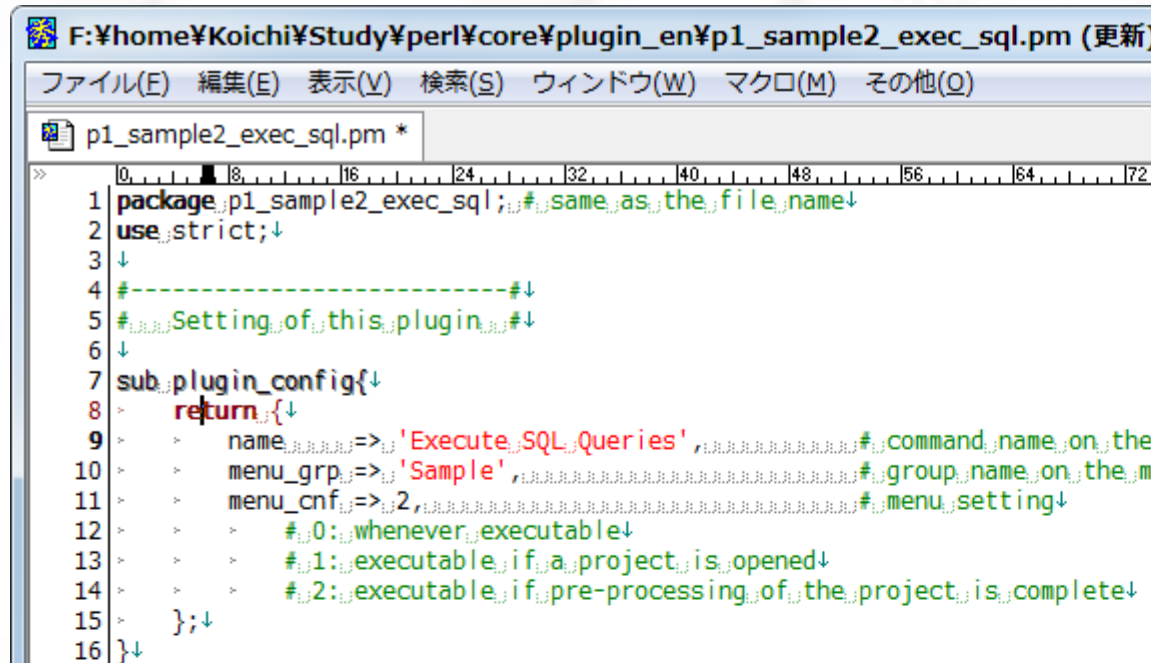
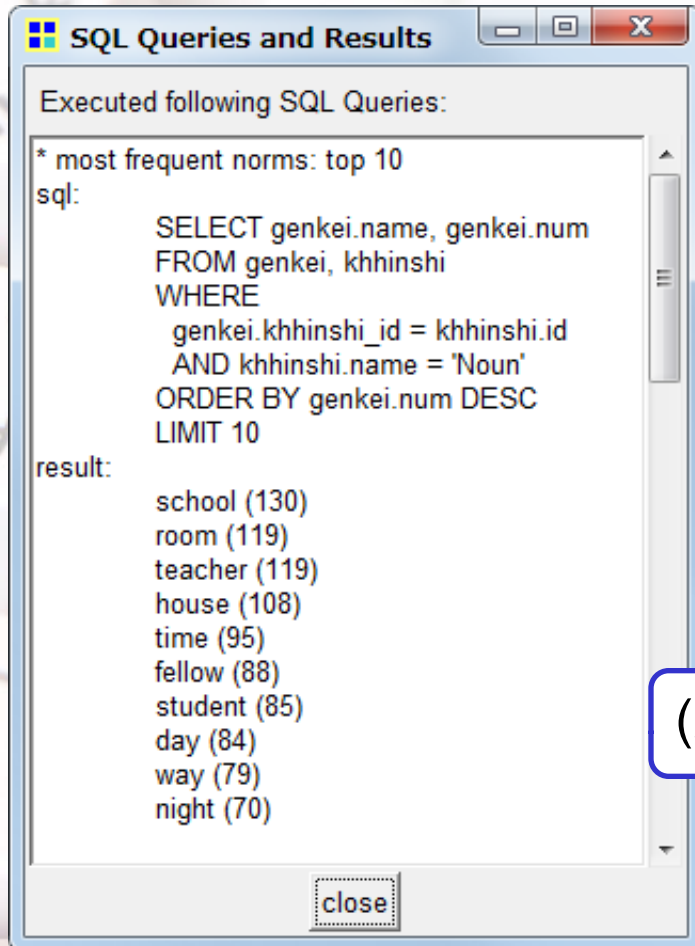
(1) Go to [Tools] [Execute SQL Statements]



- By executing SQL queries, you can bypass search functions of KH Coder, and directly utilize MySQL.
- If you start thinking GUI of KH Coder does not have enough capabilities, you can try CUI with SQL.
- For example, you cannot specify POS in "Search Words" window. But you can do it with SQL as you see.
- You can also automate the process using plugin system of KH Coder.

Sample Plugins

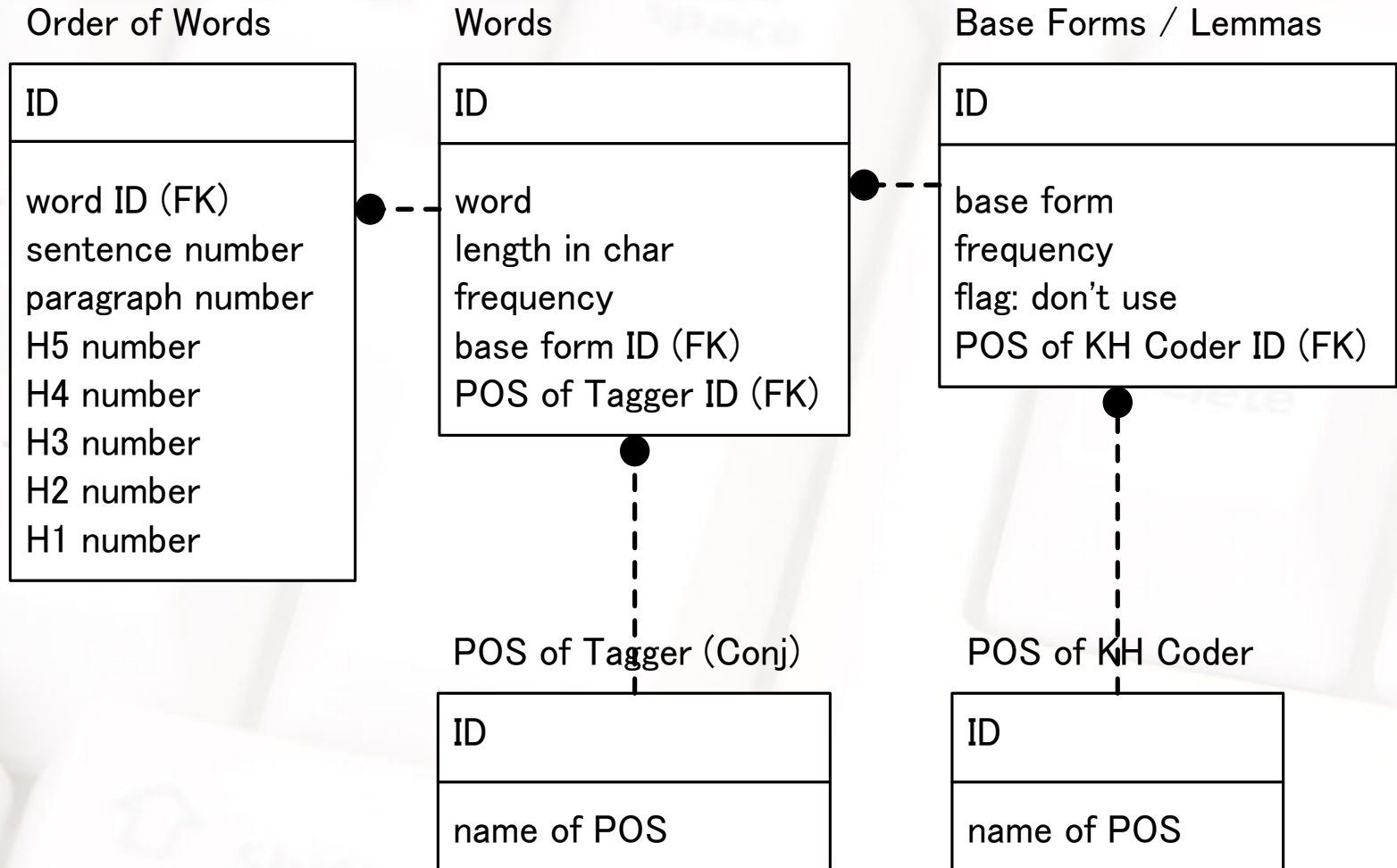
(1) Go to [Tools] [Plugin] [Sample] [Execute SQL Queries]



(2) Check "plugin_en¥p1_sample2_exe_sql.pm"

```
21 sub exec {
22     #-----#
23     # some SQL queries to run #
24     #
25     # most frequent words (Norm)
26     my $sql1 = "
27         SELECT genkei.name, genkei.num
28         FROM genkei, khhinshi
29         WHERE
```

Appendix: Tables in MySQL Database 1/2



Appendix: Tables in MySQL Database 2/2

Order of Words: “hyosobun”

Field Info	Field Name	Other
ID	Id	Key
word ID	hyoso_id	FK
sentence number	bun_id	
paragraph number	dan_id	
H5 number	h5_id	
H4 number	h4_id	
H3 number	h3_id	
H2 number	h2_id	
H1 number	h1_id	

Base Forms / Lemmas: “genkei”

Field Info	Field Name	Other
ID	Id	Key
base form	name	
frequency	num	
flag: don't use	nouse	
POS of KH Coder ID	khhinshi_id	FK

Words: “hyoso”

Field Info	Field Name	Other
ID	Id	Key
word	name	
length in char	len	
frequency	num	
base form ID	genkei_id	FK
POS of Tagger ID	hinshi_id	FK

POS of KH Coder: “khhinshi”

Field Info	Field Name	Other
ID	Id	Key
name of POS	name	

POS of Tagger: “hinshi”

Field Info	Field Name	Other
ID	Id	Key
name of POS	name	