

How Graduate Computing Students Search When Using an Unfamiliar Programming Language

Gina R. Bai
Joshua Kayani
Kathryn T. Stolee

North Carolina State University, USA

July, 2020



```
<?php
    echo "Hello, World!";
?>
```

```
disp('Hello, World!')
```

```
#include <iostream>
using namespace std;
int main()
{
    cout << "Hello, World!";
    return 0;
}
```

Hello, World!

```
console.log("Hello, World!");
```

```
print("Hello, World!")
```

```
public class HelloWorld
{
    public static void main (String[] args)
    {
        System.out.println("Hello, World!");
    }
}
```

“Hello World” in Visual Basic for Applications (VBA)

	A	B	C	D	E	F	G	H	I	J	K	L
1												
2		Goal 2: Hello World										
3												
4												
5												
6												
7												
8												
9												
10												
11												
12												
13												
14												
15												

How about This One?

Goal 4: Find Matches

Task 4: Create a button that return matched prices for certain items

4.1: Be able to return the price for one item

4.2: Be able to return prices for all items

4.3: Be able to create a button and assign this macro to it

Dictionary?

Hashtable?

HashMap?

Regex?

Sample Output			
Item	Price	Item	Price
a	\$1.00	a	\$1.00
b	\$2.00	c	\$4.00
c	\$4.00	e	\$7.00
d	\$5.00	f	\$9.00
e	\$7.00		
f	\$9.00		

Example Search Queries & Reformulations

userID: 3554248655

Graduate Student, Java & C++, Expert, Often search online

vba hashmap

vba hashmap iteration

vba dictionary iteration

vba hashtable in excel



VLOOKUP

Challenges in Code Search

Knowledge of one language can interfere with learning a new language. However, **terminology** between languages often **differs** drastically.

[Scholtz, HCI 1990]
[Shrestha, ICSE 2020]

Prior Studies on Code Search

- ❖ Professional developers & daily work
- ❖ Seek to understand
 - How and why developers search
 - Properties of search queries and search sessions
 - Code search tool
 - New tools
 - Potential improvement of existing tools

Study Design

Factors that impact the success of search queries

	Sadowski [FSE 2015]	Bai [ICPC 2020]
Duration	15 consecutive days	90-minute lab session
# Participants	27 Google developers	18 CS graduate students
Context	Daily work	5 VBA tasks
Data Types	Logs (browsing activities) & Surveys	Logs (browsing activities) & Surveys
Survey Types	1 survey, 4 questions	4 surveys, 2-6 questions each
Survey Setting	10 surveys per day 10-minute time interval	10 surveys per hour 1-minute time interval
Survey Collection	Browser extension	Browser extension

Study Design

	Sadowski [FSE 2015]	Bai [ICPC 2020]
Duration	15 consecutive days	90-minute lab session
# Participants	27 Google developers	18 CS graduate students
Context	Daily work	5 VBA tasks
Data Types	Logs (browsing activities) & Surveys	Logs (browsing activities) & Surveys
Survey Types	1 survey, 4 questions	4 surveys, 2-6 questions each
Survey Setting	10 surveys per day 10-minute time interval	10 surveys per hour 1-minute time interval
Survey Collection	Browser extension	Browser extension

Study Design

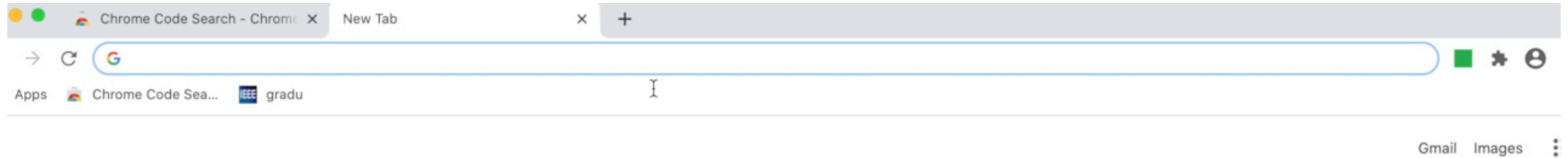
Java, JavaScript, C++, C, Python

	Sadowski [FSE 2015]	Bai [ICPC 2020]
Duration	15 consecutive days	90-minute lab session
# Participants	27 Google developers	18 CS graduate students
Context	Daily work	5 VBA tasks
Data Types	Logs (browsing activities) & Surveys	Logs (browsing activities) & Surveys
Survey Types	1 survey, 4 questions	4 surveys, 2-6 questions each
Survey Setting	10 surveys per day 10-minute time interval	10 surveys per hour 1-minute time interval
Survey Collection	Browser extension	Browser extension

Study Design

	Sadowski [FSE 2015]	Bai [ICPC 2020]
Duration	15 consecutive days	90-minute lab session
# Participants	27 Google developers	18 CS graduate students
Context	Daily work	5 VBA tasks
Data Types	Logs (browsing activities) & Surveys	Logs (browsing activities) & Surveys
Survey Types	1 survey, 4 questions	4 surveys, 2-6 questions each
Survey Setting	10 surveys per day 10-minute time interval	10 surveys per hour 1-minute time interval
Survey Collection	Browser extension	Browser extension

Chrome Extension Demo – Logs & Surveys



Time Interval between Surveys / Clicks

#Query, #Click, #avgClick/Query

Types of Visited Websites

Query Responses

Overview of Logged Data



18 participants



3,508 log entries



229 queries

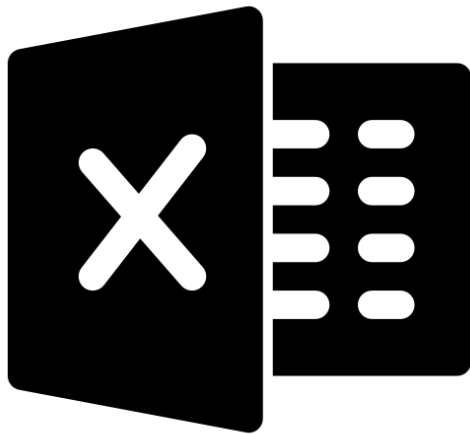


216/389 deployed surveys were completed

Research Questions

- ❖ RQ1 – Why do subsequent language learners search?
- ❖ RQ2 – What does a typical search session entail?
- ❖ RQ3 – What are the factors that impact the success of search queries?

RQ1 – Why do learners search?



Example Code



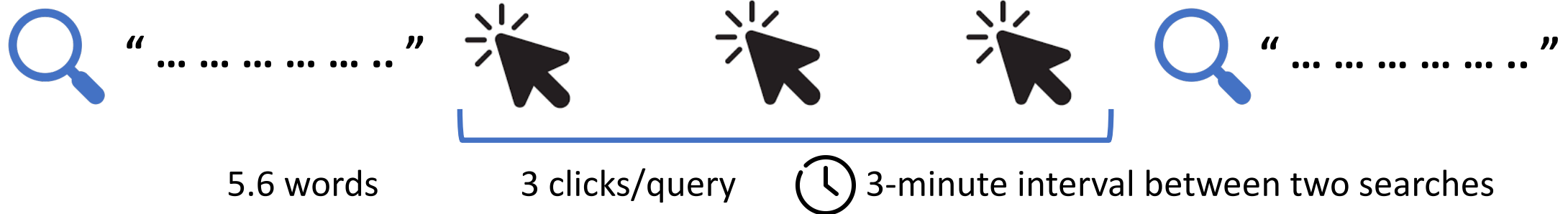
Resolve Bugs/Errors



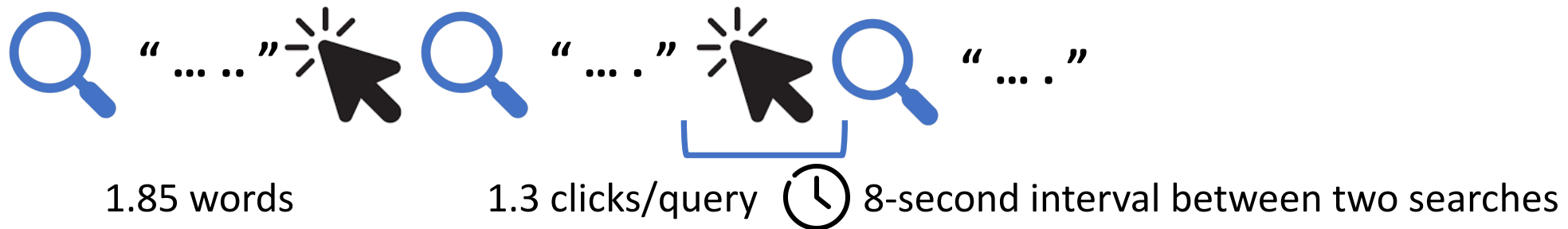
All searched
for exploring

RQ2 – Typical Search Session

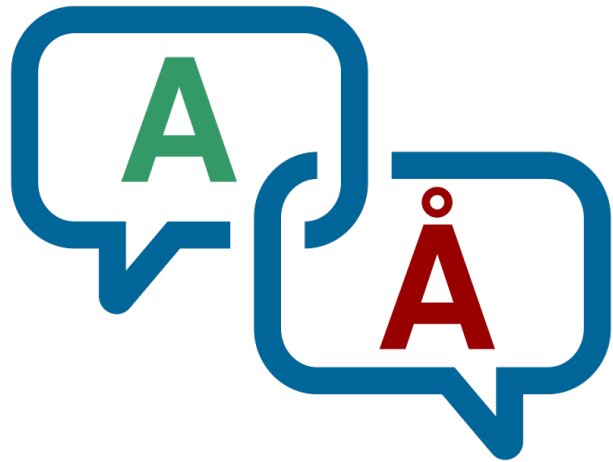
Subsequent language learners



Professionals Sadowski [FSE 2015]



RQ3 – Success of Search Queries



Borrow terms from familiar languages



Consult tutorials & documentation rather than Q&A



Use natural languages & Consult others

Takeaways

- ❖ Reformulations occur much more quickly after a previous query than searching for a new topic
- ❖ The word-based Levenshtein distance between successive queries alone may not indicate a reformulation
- ❖ Term borrowing was commonly observed
- ❖ The impact of term mismatch in terms of query success and reformulation is fairly neutral