



Sascha Wolfer, Martina Nied, Idalete Dias, Carolin Müller-Spitzer, Majo Dominguez

COMBINING QUANTITATIVE AND QUALITATIVE METHODS IN A STUDY ON DICTIONARY USE

EURALEX Ljubljana
July 17th, 2018

OUTLINE

- **General remarks**
- **Study design**
- **Qualitative & Quantitative Analyses:**
 - Factors which contribute to a successful search and the accuracy of the corrections**
 - Time participants spend on the resources
 - Language awareness, dictionary use awareness, determination
 - Number of the consulted resources
 - Type of the consulted resources
 - Intuition and focalization during the search process
- **Conclusions**

STARTING POINT

- Modern lexicography recognizes and benefits from dictionary usage research.
- Numerous and often sophisticated studies about
 - what users appreciate about dictionaries
 - which information they need in different situations
 - how and how fast they find specific information
 - ...
- Presupposition: People with a language problem **do consult dictionaries.**

MOTIVATION

- In contrast, language teachers often have the impression that students use **too few** high-quality lexicographic resources in their everyday work.

Given that, many studies might start at a point that many students never reach when they are dealing with everyday language problems.

MOTIVATION

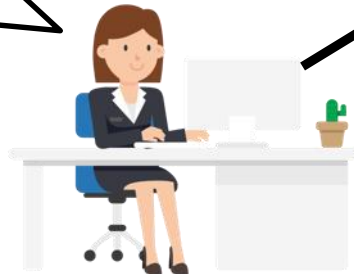
- Overarching research question: What do students actually **do** when they try to correct language problems?
- Current focus: L2 students of German with a Romance native language (Italian, Portuguese, Spanish).

Multi-methodological exploratory observational study with an explicit focus on the interplay between quantitative and qualitative analyses

SETUP

- **43 participants** from Braga (Portugal), Santiago de Compostela (Spain) and Rome (Italy), CEFR level of German: A2/B1.
- **Task:** Correct 18 German sentences with interference errors (45 minutes per participant).
- **Observation:** Screen & voice recordings

allora cerco ehm vivere leben
vedo se mi da qualche
informazione # ok cerco

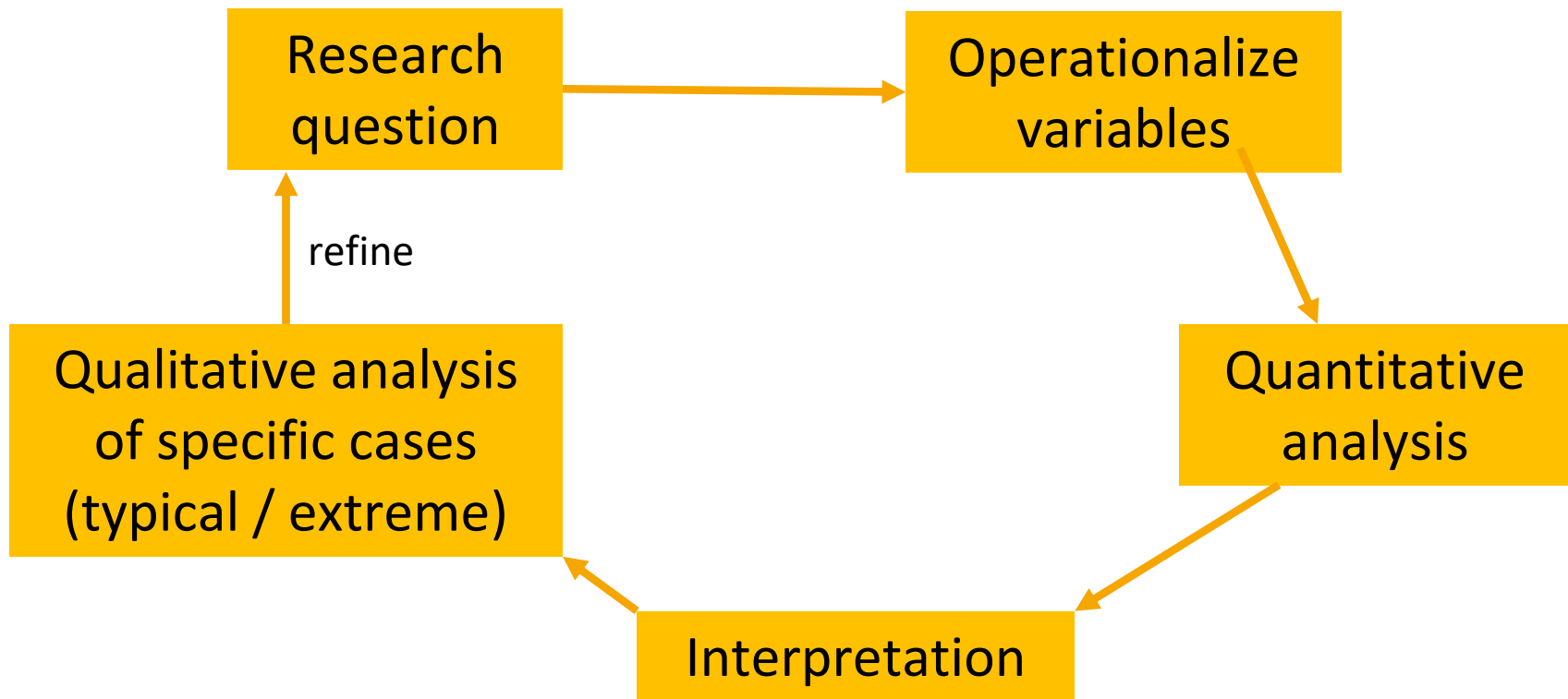


Satz	Korrektur
Morgen habe ich einen Arzttermin und kann deshalb nicht in die Klasse gehen.	
Der Artikel handelt sich um die Migranten in Deutschland.	
Bist du bereit? Wir müssen jetzt los, wir sind sowieso schon zu spät dran.	
Meine Nachbarin möchte immer alles wissen. Sie ist sehr kurios.	
Ich bin einverstanden mit dir.	
Ich möchte dir heute über einen interessanten Artikel sprechen.	
Obwohl ich studiere, wohne ich noch mit meinen Eltern.	
Wenn ich zur Schule ging, habe ich viel Sport gemacht.	
An unserem Forschungsinstitut ist Ihnen unsere Bibliothek 24 Stunden zur Verfügung.	
Obwohl sich der Junge beeilt hat, hat er die U-Bahn verloren.	
Ich vorbereite gerade meine letzte Prüfung.	
Kein Problem, wenn der Zucker beendet ist; ich nehme dann Honig.	
Ich habe die Hose viel zu klein gekauft. Jetzt muss ich nochmals ins Geschäft zurück und sie wechseln.	
Ich möchte ein Stipendium beim DAAD bewerben.	
Er wohnt seit Jahren in Berlin und trotzdem verliert er sich immer noch.	
Leider kann ich heute nicht Tennis spielen. Ich bin zu besetzt.	
Um beim Kartenspielen zu gewinnen, musst du exakt die Regeln folgen.	
Das erlaube ich dir nicht. Es ist außer Frage.	

DATA PREPARATION

- Annotate 671 correction results.
- Transform approx. 32 hours of screen recordings into approx. 7500 “actions”.
- Transcribe approx. 32 hours of voice recordings.
- Categorize 2235 search strings.
- Categorize 60 used resources.
- Integrate all data sources into one large data structure (cf. our full paper for an example).

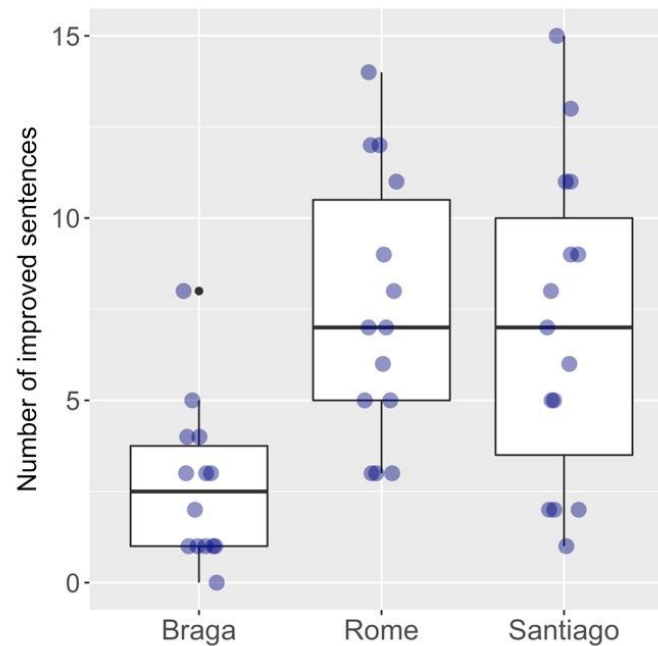
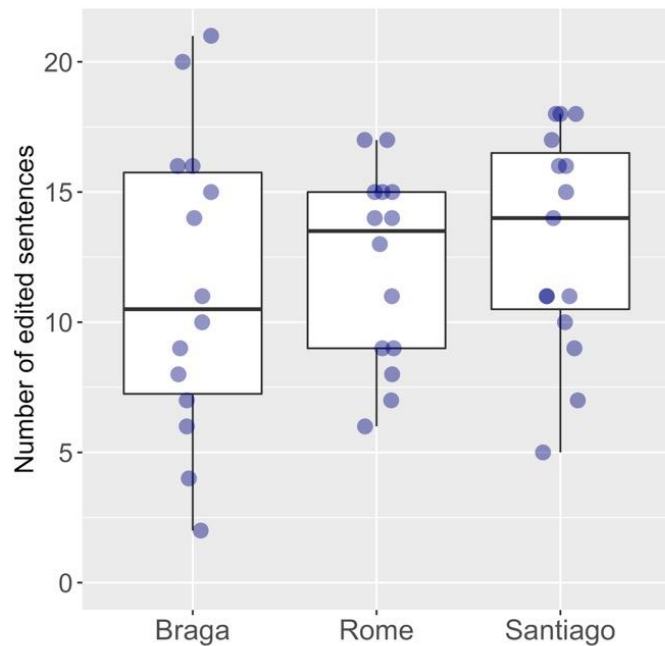
COMBINATION OF QUANTITATIVE AND QUALITATIVE RESEARCH



OUTLINE

- **General remarks**
- **Study design**
- **Qualitative & Quantitative Analyses:**
 - Factors which contribute to a successful search and the accuracy of the corrections**
 - Time the participants spend on the resources
 - Language awareness, dictionary use awareness, determination
 - Number of the consulted resources
 - Type of the consulted resources
 - Intuition and focalization during the search process
- **Conclusions**

CORRECTED/CORRECTLY IMPROVED SENTENCES IN THE THREE LOCATIONS

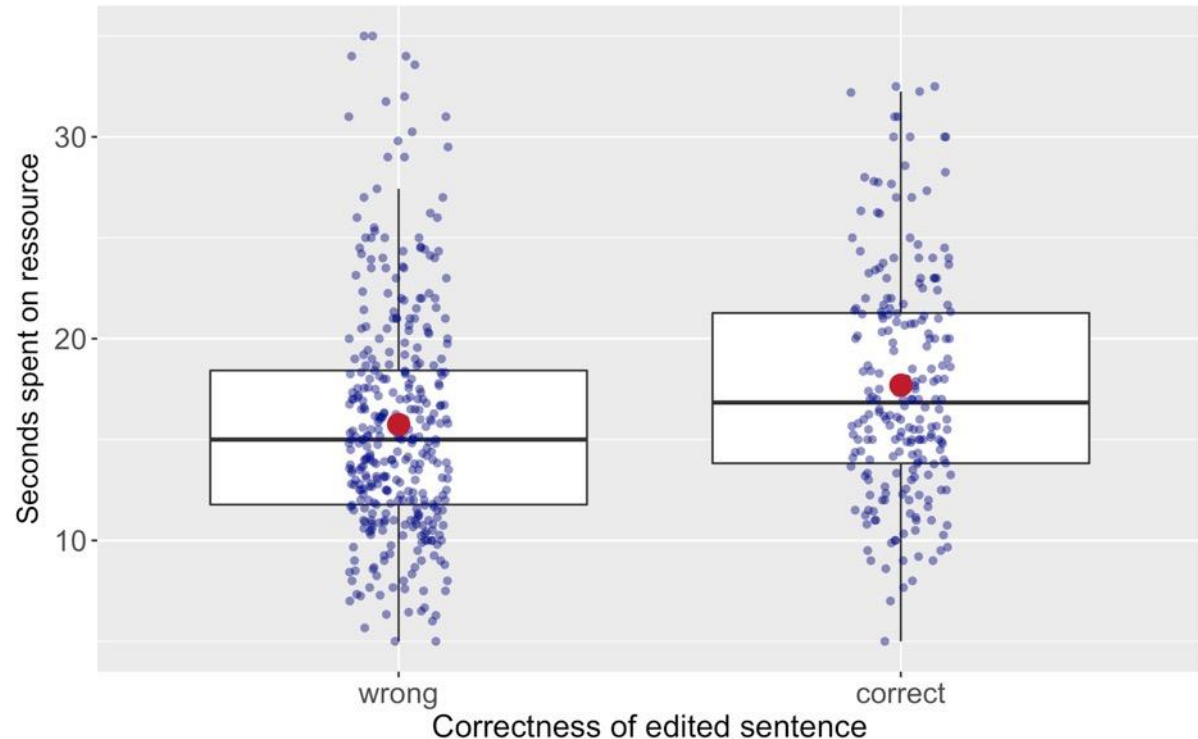


- No. of edited sentences very similar across universities
- Significant difference in the percentage of correct sentences across universities

TIME SPENT ON RESOURCES AND CORRECTNESS (QUANTITATIVE ANALYSIS)

Question: Is there a relation between time participants spend on resources and quality of the corrections?

- Average difference in seconds is small, but significant:
- The time spent on each resource is 2.4 seconds longer in edits that result in a correct sentence.



TIME SPENT ON RESOURCES AND CORRECTNESS (QUALITATIVE ANALYSIS)

“ZAPPER MENTALITY”

Questions:

1. The less time students spend with a single resource and the faster the backward and forward switching between resources, the more difficult it becomes for them to gain a clear idea of how to solve the problem?
2. Or the less motivated they become to really solve the problem?
3. Is there a relation between the forward and backward switching and not solving the exercise?

Timestamp	Name Resource	Intuition before search? (j n u)
00:00:39	Excel	n
00:01:33	Excel	n
00:04:08	Google.it	n
00:04:26	Dicios.it	n
00:04:28	Excel	n
00:04:30	Dicios.it	n
00:04:40	Google.it	n
00:04:44	Excel	n
00:04:47	Google.it	n
00:04:54	Google.it	n
00:04:57	LEO Dictionary	n
00:05:10	Excel	n
00:05:14	Google.it	n
00:05:30	Deutsch Info	n
00:05:38	LEO Dictionary	n
00:05:46	Excel	n
00:05:58	Excel	n
00:06:50	LEO Dictionary	n
00:06:56	Excel	n
00:07:02	LEO Dictionary	n
00:07:04	Deutsch Info	n
00:07:07	Google.it	n
00:07:24	Deutsches Institut Florenz.it	n
00:08:00	Excel	n
00:08:09	Deutsches Institut Florenz.it	n
00:08:13	Excel	n

TIME SPENT ON RESOURCES AND CORRECTNESS (QUANTITATIVE ANALYSIS)

“ZAPPER MENTALITY”

R 02: Ø 10 sec. per resource/Excel-Spreadsheet

R 03: Ø 21,6 sec. per resource/Excel-Spreadsheet

Sentence	Time (sec.)	Switch	Ø Sec./Resource	Correct
1	391	40	10	N
2	93	12	7,75	N
3	367	39	9,4	N
4	125	10	12,5	Y
5	189	25	7,5	N
6	159	14	11	Y

Sentence	Time (sec.)	Switch	Ø Sec./Resource	Correct
1	189	8	23	Y
2	88	5	18	Y
3	169	9	18	Y
4	234	7	33	Y
5	300	20	17	Y
6	252	8	31	N

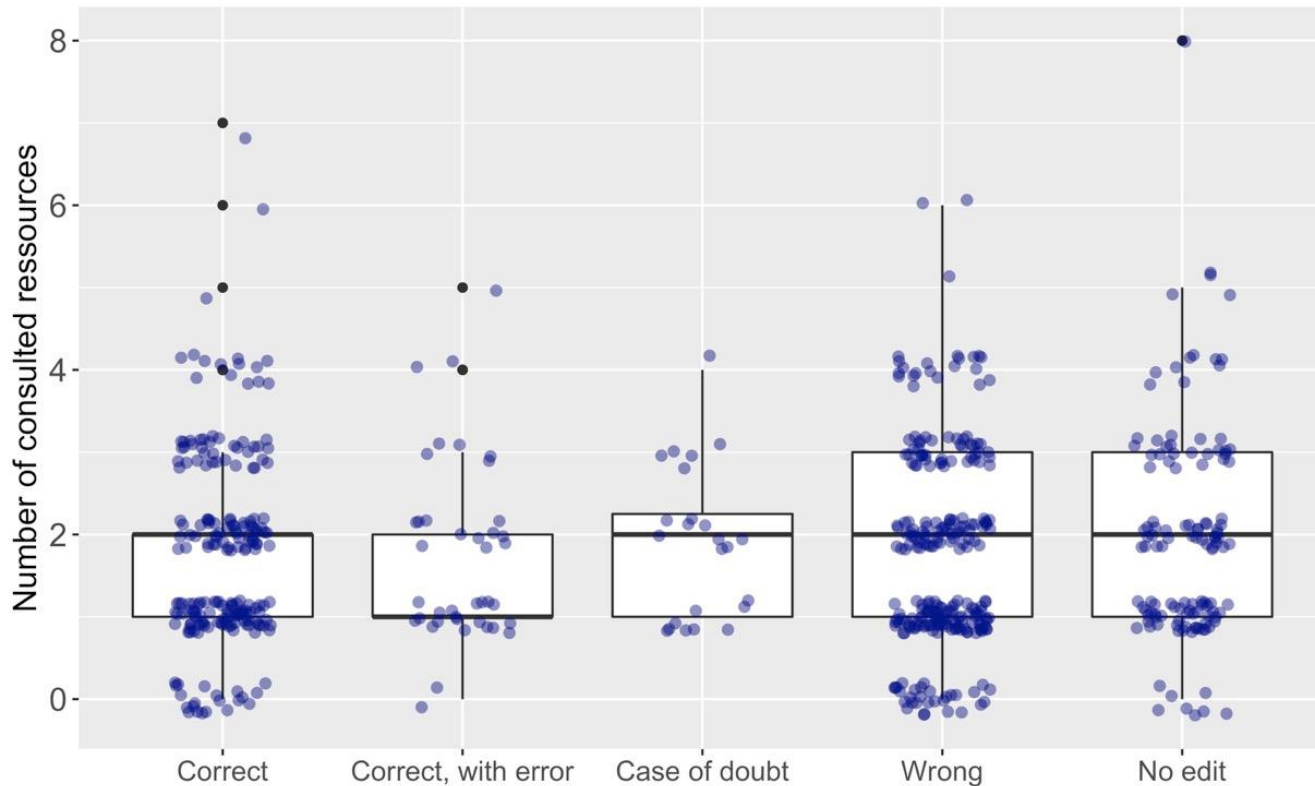
- The less forward and backward switching actions between resources, the better the result?!
- The more time spent with a specific resource, the better the result?!

LANGUAGE & DICTIONARY USE AWARENESS (QUALITATIVE ANALYSIS)

R-07 *Obwohl sich der Junge beeilt hat, hat er die U-Bahn **verloren**.*
*“Although the boy hurried, he **lost** the subway.”*

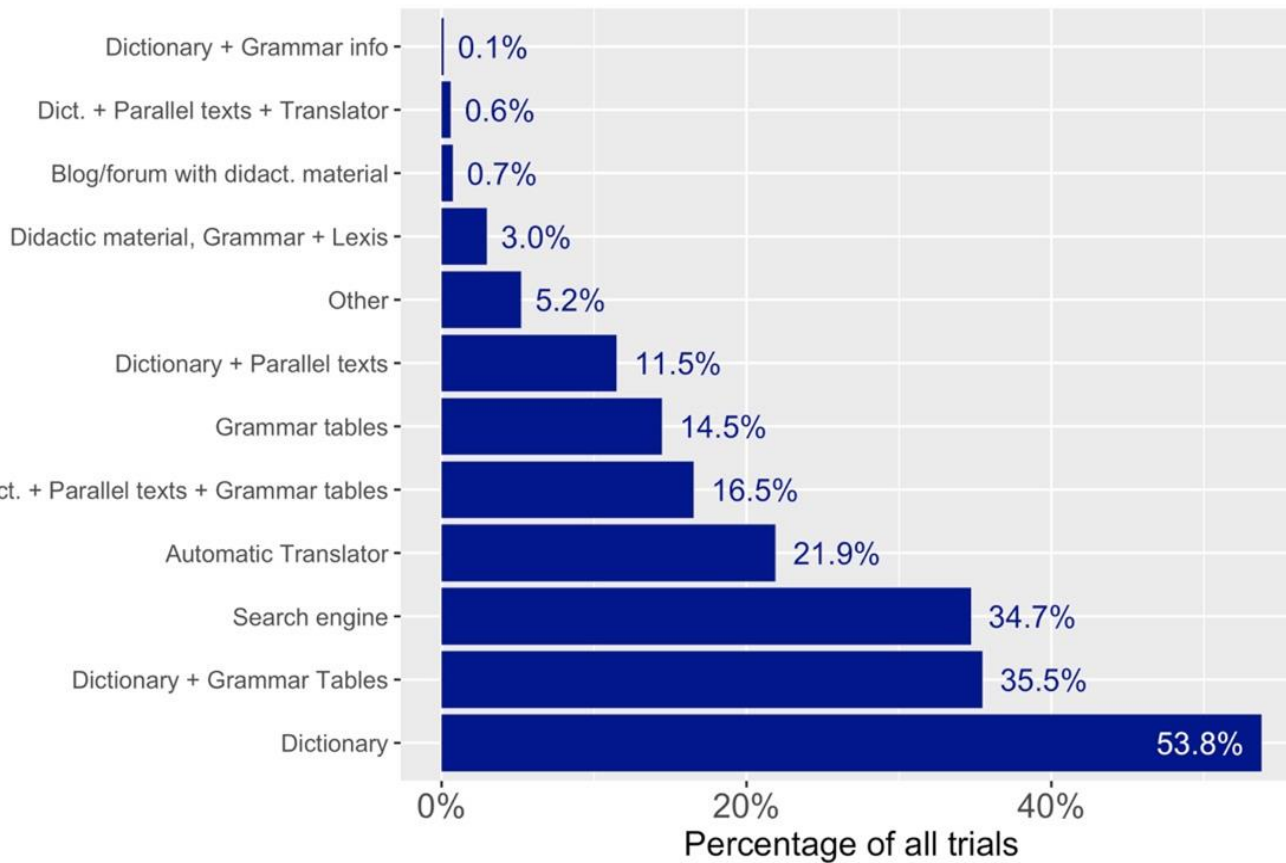
1 So I read the first sentence [she reads the sentence] (Obwohl sich der Junge beeilt hat, hat er die U-
 2 Bahn verloren) eh the section in bold is verloren so the verb ehm there are two sentences therefore
 3 one is the main clause and one is the subordinate clause ehm so first I look for the verb verloren I
 4 prefer google chrome # ehm # generally I use the online dictionary PONS because perhaps it shows
 5 also the context and the use of a word and also some examples so therefore I search from German to
 Italian okay
 6 so verloren ## which means # okay I see it's the past participle of the verb verlieren so lost the past
 7 participle 'perduto' lost ok # so I lost the train even though the boy (hat beeilt) I look it up because I'm
 8 not sure what does it mean so # beeilt (affrettarsi) okay (sich beeilen) so even though the boy hurried
 9 he lost the train #perhaps the error would be that verloren is used in other contexts so I look for other
 10 use contexts or a synonym of the verb verlieren so # I look up verlieren and it shows me (perdere
 11 perdere la testa # disperdersi) okay
 12 In Italian I look for a synonym of the verb perdere in German so I set Italian German and look for
 13 perdere okay so (verlieren verlegen smarrire) # eh perdere il treno for example it shows (verpassen)
 14 which means that I can use verpassen instead of verlieren and ehm okay it shows me (Zug) so okay I
 15 use verpassen I will look for the past participle for being sure so I open a website with the name
 16 **REVERSO**
 17 ## so conjugation German verbs I search on google and it gives me REVERSO so I look for verpassen #
 18 so (verpasst) okay I write again the sentence on the right obwohl sich der Junge beeilt hat hat er die U-
 19 Bahn verpasst [she finishes the sentence] okay I go continue with the second sentence

NUMBER OF CONSULTED RESOURCES AND CORRECTNESS



The number of resources used has no influence on the correctness of the sentence

CATEGORIZATION OF RESOURCES



- Types of resources used for all edits
- Large number of different types of resources used
- Dictionaries are used most often, followed by dictionaries with grammar tables

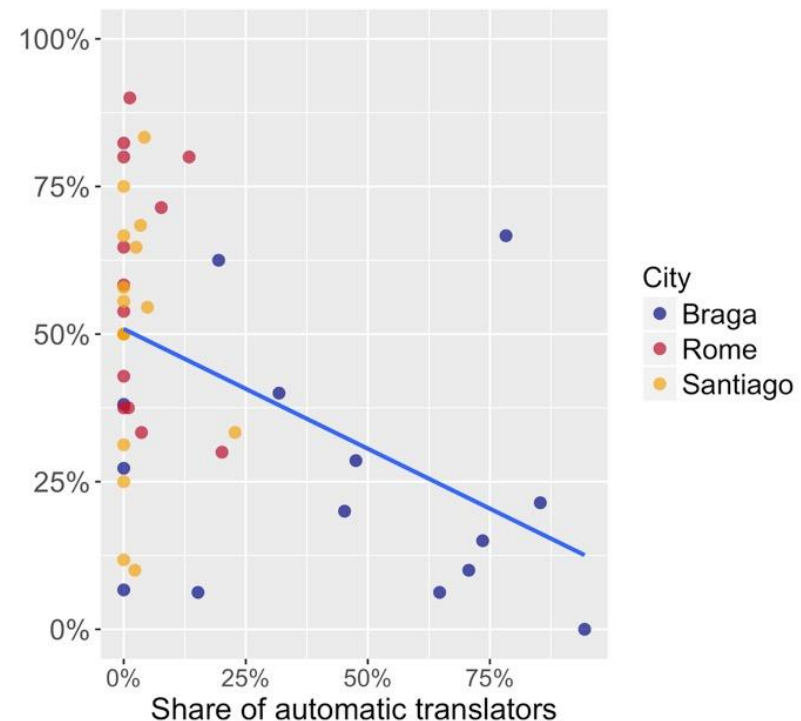
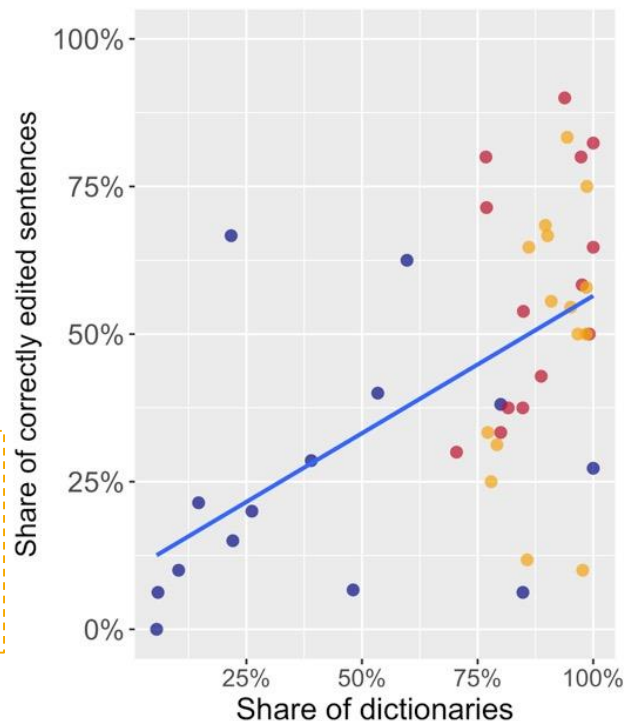
TYPES OF RESOURCES AND CORRECTNESS

Participants who used more **dictionary** resources were **more successful**.

Participants who rely more on **automatic translation programs** have achieved **poorer revision results**.

Dictionaries with parallel texts tend to produce even better results.

This connection must be examined in more detail.



INTUITION AND FOCALIZATION HYPOTHESIS

Situation 1:

E.g.: *Er wohnt seit Jahren in Berlin und trotzdem **verliert** er **sich** immer noch.*

- 8 students assume (intuition before initiating the search process/ hypotheses resulting from search actions) that the verb ‘verlieren’ is correct in this context.
- **All further search actions start out from this assumption:**
 - ‘verlieren’ / ‘sich verlieren’ is correct;
 - The problem lies in the reflexive pronoun: ‘verlieren’ does not take a reflexive pronoun;
 - The problem lies in the word order of the verb ‘verlieren’ or the word order of the reflexive pronoun.
 - **St. try to confirm/validate the hypotheses.**
- Search actions that take ‘sich verlieren’ as a starting point (*Linguee, Pons, Google Translate*), lead to unsuccessful results.

INTUITION AND FOCALIZATION HYPOTHESIS

E.g.: *Er wohnt seit Jahren in Berlin und trotzdem **verliert** er **sich** immer noch.*

[B-08]

Search action (1): St. enters 'verliert' in GOOGLE-TRANSLATE.

- **TAP:** *Google-Translate to understand the meaning of the sentence*
- **TAP:** *(**verliert**) # ahhh ele mora anos lá e mesmo assim **perde-se** #*
- **TAP:** *er wohnt seit Jahren in Berlin und trotzdem verliert er sich dort (verliert er sich # sich verliert) # mmm I don't know if it is ,sich verliert' # ,verliert sich')*

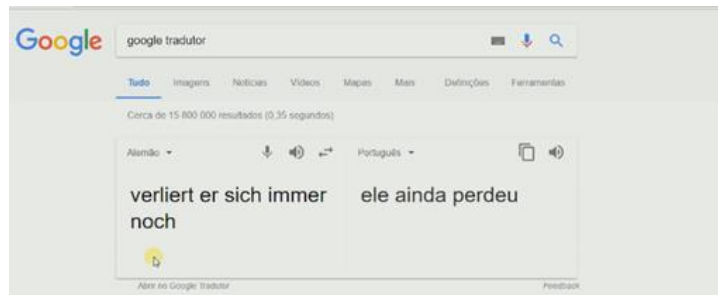


Search action (2): St. enters 'verliert sich' in GOOGLE-TRANSLATE.



Search action (3): St. enters 'verliert er sich immer noch' in GOOGLE-TRANSLATE.

- **TAP:** *(er wohnt seit Jahren in Berlin und trotzdem verliert er immer noch) # I think that (**sich**) is not needed here #*



Search action (4): Changes language direction: PT – DE

INTUITION AND FOCALIZATION HYPOTHESIS

E.g.: *Er wohnt seit Jahren in Berlin und trotzdem **verliert er sich** immer noch.*

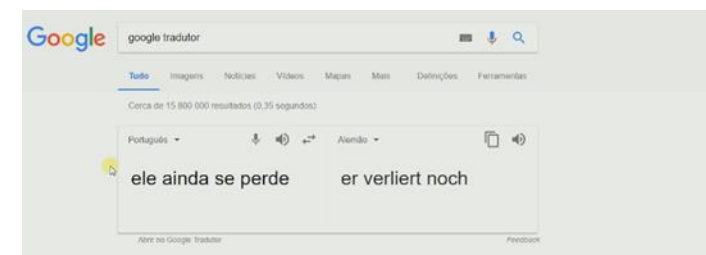
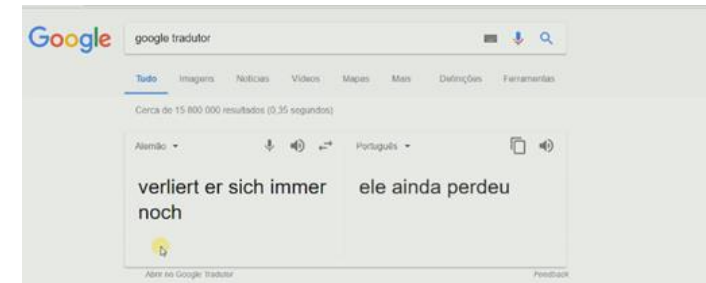
Search action (4): Changes language direction: PT – DE

[The pt. equivalent ‘perde-se’/ ‘perder-se’ does not appear in the GOOGLE-Translate result »» this does not correspond to the st. expectation.

St. expects ‘perde-se’. Changes language direction: ‘perder-se’ – ‘sich verlieren’ or just ‘verlieren’ (without ‘sich’)]

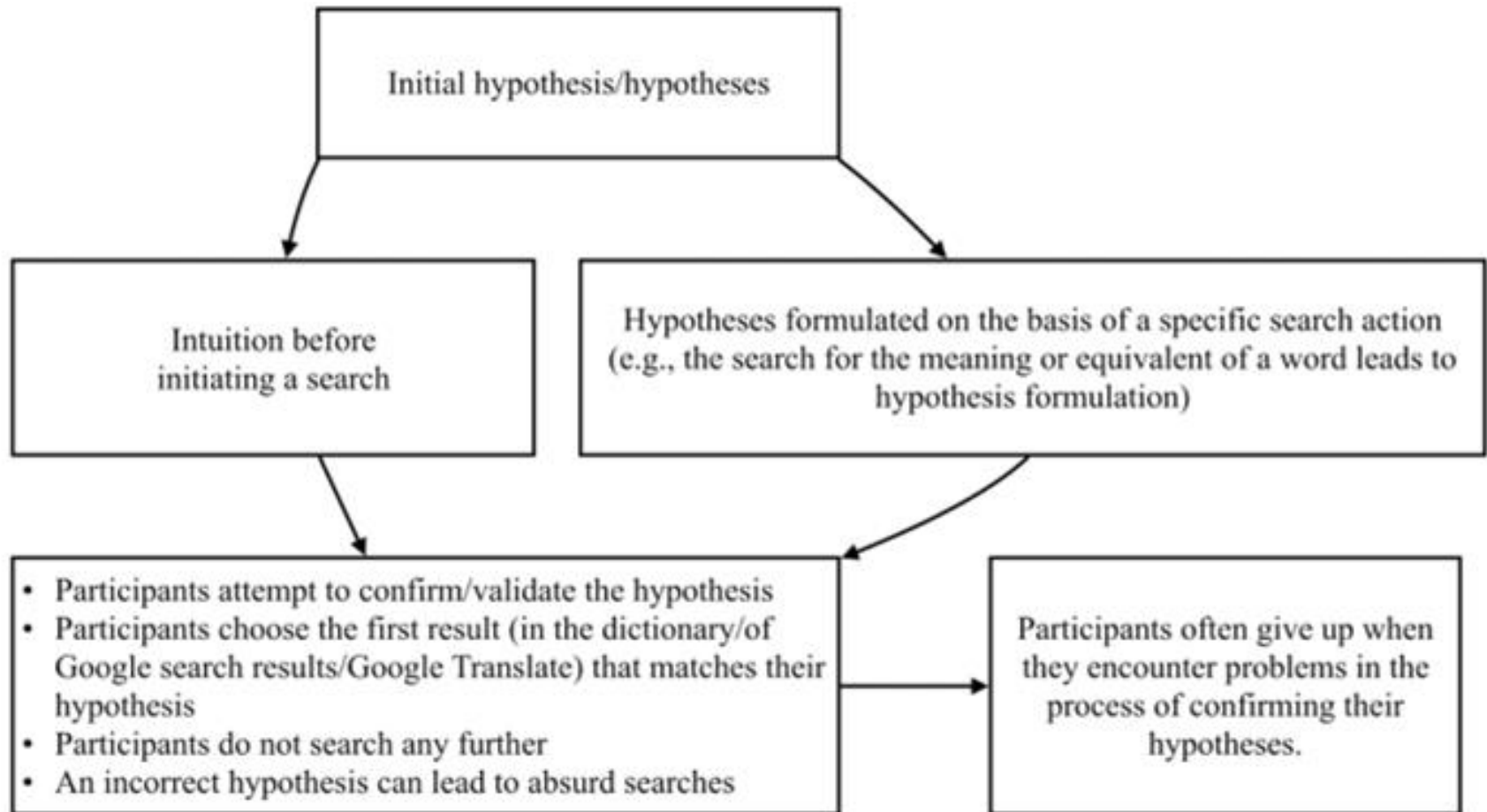
- St. enters ‘ele ainda **se** perde’.
- German GOOGLE equivalent ‘er verliert noch’ – without ‘sich’
 - **TAP:** (*er wohnt seit Jahren in Berlin und trotzdem verliert er immer noch*) # **exactly I think it doesn't need (sich)** #

Correction: *Er wohnt seit Jahren in Berlin und trotzdem verliert er immer noch.*



A good example showing that false hypotheses can lead to absurd searches. [Focalization Hypothesis!]

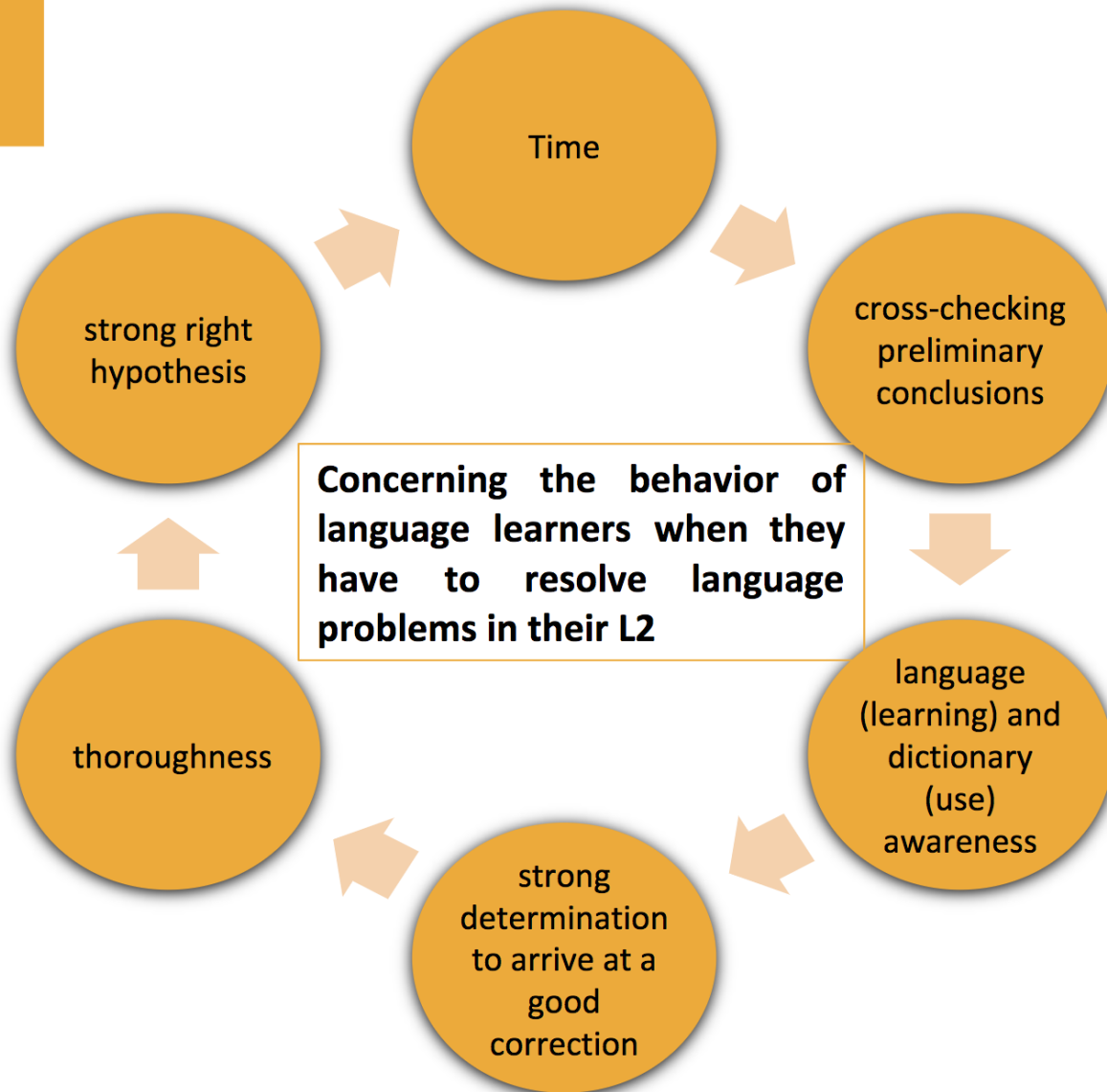
INTUITION AND FOCALIZATION HYPOTHESIS



OUTLINE

- **General remarks**
- **Study design**
- **Qualitative & Quantitative Analyses:**
Factors which contribute to a successful search and the accuracy of the corrections
 - Time the participants spend on the resources
 - Language awareness, dictionary use awareness, determination
 - Number of the consulted resources
 - Type of the consulted resources
 - Intuition and focalization during the search process
- **Conclusions**

CONCLUSIONS



Some factors (strong determination, good metalinguistic knowledge etc.) are very hard to operationalize on a quantitative level

we presented evidence from the verbal protocols

CONCLUSIONS

- Iterative process: quantitative results – qualitative analyses – quantitative analyses
- Interesting and relevant data for research into dictionary use
- Closer to everyday reality of DaF learners, i.e. closer to the reason why language resources are consulted
- Small sample of informants, but on the other hand a considerable amount of data from each participant
- Data processing is costly and time-consuming
- Exploratory study, i.e. it is not meant to answer very specific questions
- What would we do differently in the future?
 - Apply a language level test beforehand; have students perform a translation task

THANK YOU!



Martina Nied Curcio

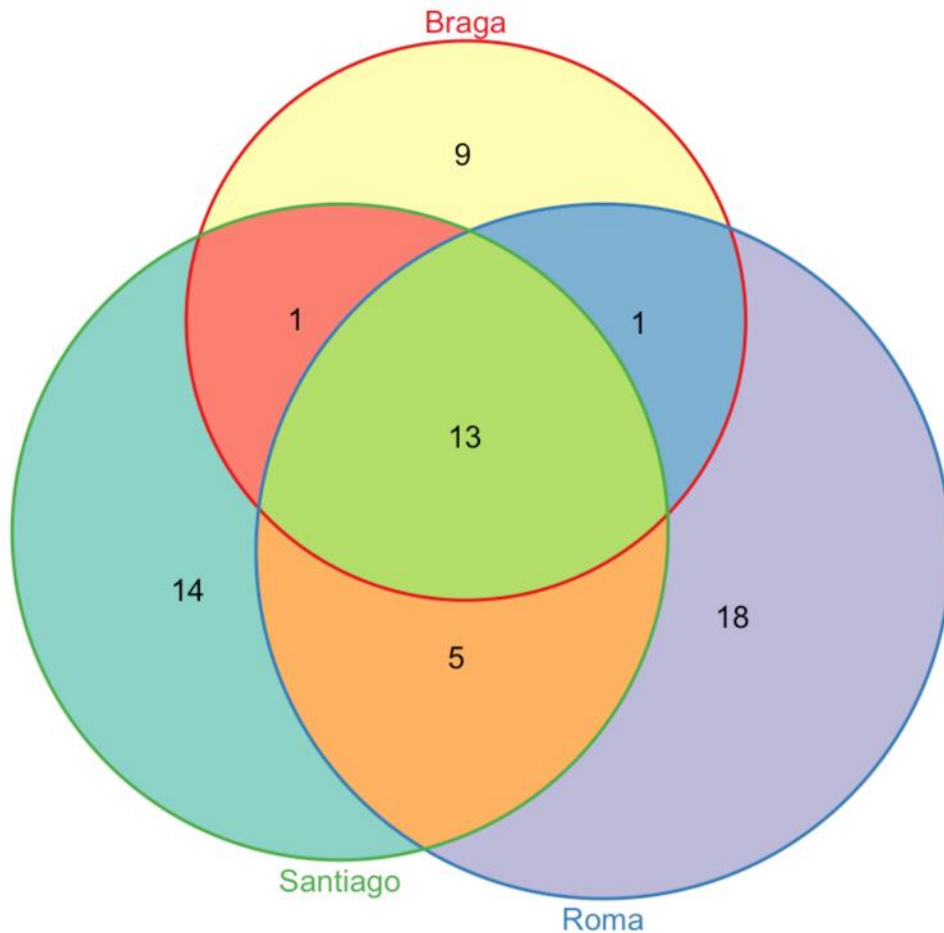


Carolin Müller-Spitzer

Submitted: More extensive article for Lexikos:

“Right hypotheses and careful reading are essential: Results of an observational study on learners using language resources online.”

BONUS SLIDE: WHICH RESOURCES WERE USED?



Resource	% Participants	% Edits
Google	100,0%	38,6%
Pons Dictionary	79,1%	54,6%
Pons Translator	37,2%	11,0%
Reverso Conjugation	34,9%	7,29%
Linguee	32,6%	11,0%
Google Translator	27,9%	16,3%
Verbformen	27,9%	7,12%
Duden	25,6%	7,12%
Bab,la Dictionary	16,3%	5,59%
Dict.cc	14,0%	1,19%
Deutsch Info	14,0%	3,39%
Lingolia Deutsch	11,6%	1,86%
WordReference	9,3%	2,71%

Large number of resources used = **60**

BONUS SLIDE: VISUALISATION OF EDITING PROCESS

