



20 results ▾

SEARCH 🔍

CONFIGURE ⚙️

READ ☰

# Automatic Input Enrichment for Selecting Reading Material

## Online Study with English Teachers

**Maria Chinkina, Ankita Oswal, Detmar Meurers**  
**BEA13 in New Orleans, USA**  
**06/05/2018**



---

# Outline

1. Motivation
2. Input enrichment
3. FLAIR system
4. The online study
5. Discussion and Outlook



## Motivation

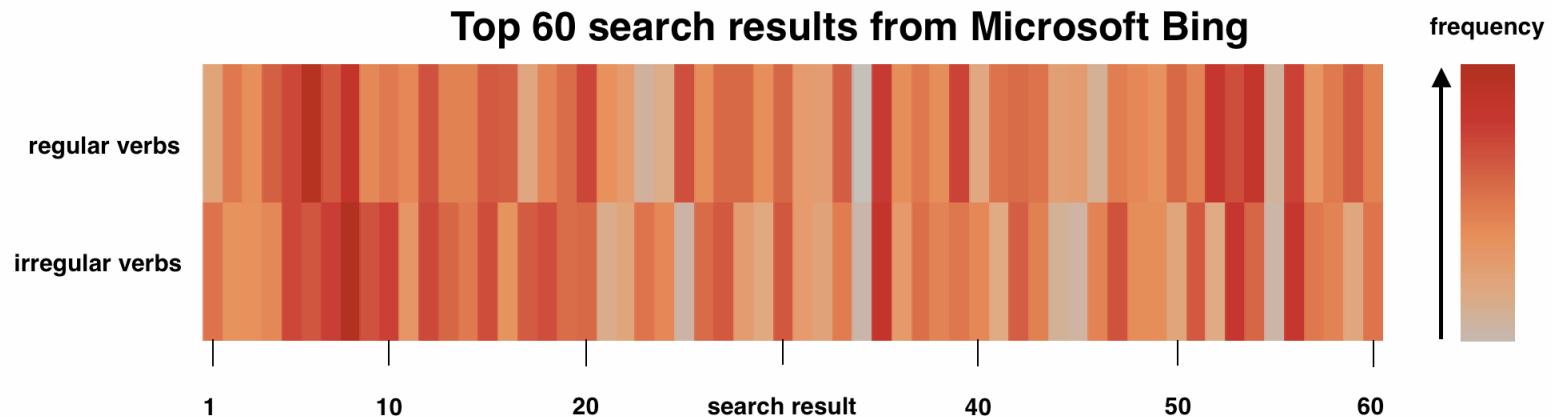
- **Input Hypothesis** (Krashen, 1977)
  - Exposure to input containing target constructions
- **Frequency of forms** (Slobin, 1985)
  - The more frequent the construction in input, the more it facilitates acquisition
- **Input Enrichment**  
(Related to *input flood* by Trahey and White, 1993)
  - Ensuring a high number of target constructions in text



## Problem

Material of interest is readily available **on the Web**.  
But the linguistic forms are **unevenly distributed** across texts.

The top 60 results for the query *Brexit*:

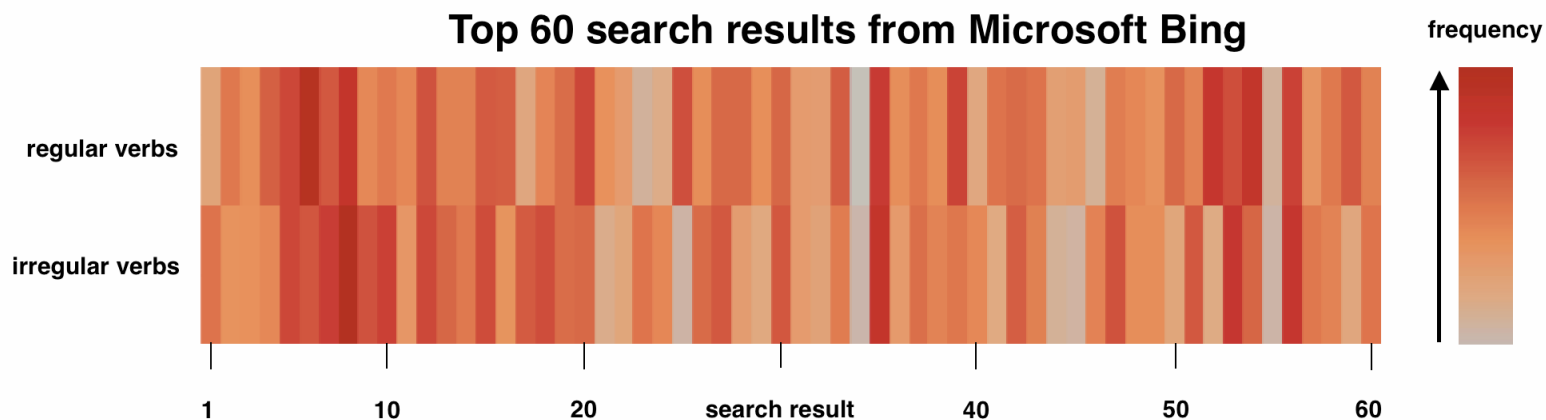




## Problem

Material of interest is readily available **on the Web**.  
But the linguistic forms are **unevenly distributed** across texts.

The top 60 results for the query *Brexit*:

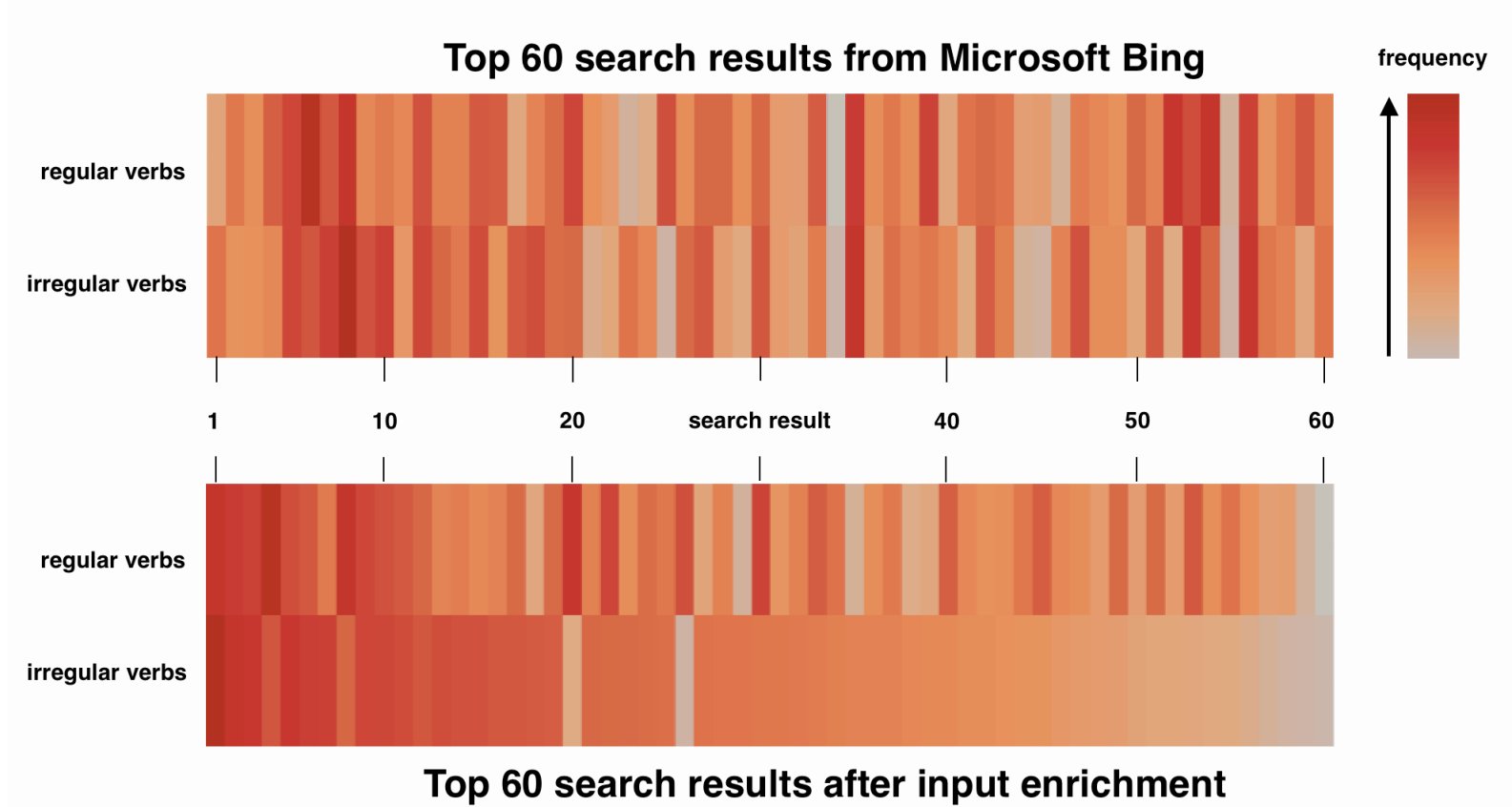


1. How do teachers and learners get to the best results?
2. Are those results still good content-wise?



## Solution: Automatic input enrichment

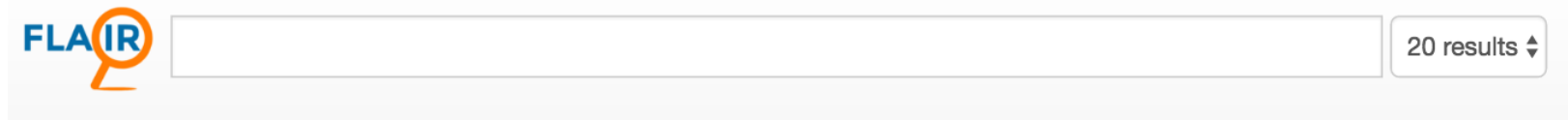
We re-rank the search results based on the representation of the selected linguistic forms in them.





## FLAIR: Web Search for Language Teachers

Available online: [purl.org/ical1/flair](http://purl.org/ical1/flair)



**FLAIR** (Chinkina & Meurers, 2016) helps teachers search for texts appropriate in **form**, **content**, and **reading level**. It:

- **retrieves** the search results from Microsoft Bing or lets the user **upload** their own texts,
- **identifies** 87 linguistic forms from the official curriculum,
- **reranks** the results based on the selected linguistic forms,
- thus, **provides** systematic automatic input enrichment.

# FLAIR Interface

← FLAIR
🕒 🌐

10 Results (0 Filtered)

VISUALIZE ↗

SHARE SEARCH SETUP ↩

Text Characteristics: ☰

Length:

Prefer shorter texts.

Levels:

✓ A1-A2 6/10

✓ B1-B2 7/10

✓ C1-C2 7/10

Constructions: ✓

▼ Sentences ↶

▼ Parts of Speech ↶

Language Use: 📄

Academic Vocabulary ✍ ↶

RESET ALL ↶

'Brexit site:reuters.com'

- 1 ⋮

[Brexit group fined for breaking spending rules in EU vote ...](#)  
https://www.reuters.com/article/uk-britain-eu-spending/britains...

Britain's Electoral Commission imposed a record-matching 70,000 pound fine on Friday on one of the main groups that campaigned for Brexit, and said police might have to investigate possible criminal offences because of unreported campaign spending.
- 2 ⋮

[Exclusive - Jumping ship: Brexit-hit EU staff ditch UK ...](#)  
https://www.reuters.com/article/us-britain-eu-jobs-exclusive/...

About one in 10 British civil servants at the European Commission has taken another EU nationality since the Brexit vote, but are nonetheless resigned to scant prospects of future promotion.
- 3 ⋮

[Brexit seen threatening UK links in EU supply chain | Reuters](#)  
https://www.reuters.com/article/uk-britain-eu-aviation/brexit-seen...

Mandy Ridyard knew Brexit was going to be a challenge for her aviation components firm, but it was still a shock when she heard a French company bluntly ruling out British suppliers from an international bid for a contract in China.
- 4 ⋮

[UK's May facing local election losses, Brexit unity at ...](#)  
https://www.reuters.com/article/us-britain-politics-election/uks...

Voters in England cast their ballots on Thursday in local government elections expected to show rising support for Prime Minister Theresa May's opponents in London and add to questions about her ability to follow through on her Brexit plan.
- 5 ⋮

[Whistleblower questions Brexit result, says campaigners ...](#)  
https://www.reuters.com/article/uk-britain-eu-lawyers/...

A whistleblower at the heart of a Facebook data scandal on Monday questioned the result of Britain's 2016 Brexit referendum as his lawyers presented evidence that they said showed the main campaign for leaving the EU had broken the law.
- 6 ⋮

[Trust me on Brexit, May says as ministers squabble | Reuters](#)  
https://www.reuters.com/article/uk-britain-eu/trust-me-on-brexit...

British Prime Minister Theresa May said on Sunday she could be trusted to deliver Brexit, but that it could not be done without compromises on all sides – a possible warning to cabinet ministers who are

Brexit group fined for breaking spendi... ✕

B1-B2 34 Sentence(s) 631 Word(s)

referendum. The commission **said** it **was** enforcing the rules without suggesting the breach **had altered** the result. "These are serious offences," **said** Bob Posner, the Electoral Commission's director of political finance and regulation. "Leave.EU **exceeded** its spending limit and **failed** to declare its funding and its spending correctly". **Asked** on BBC radio if the commission **was** saying the breach **was** serious enough to have **impacted** the result, its chief executive, Claire Bassett, **said** "No", but she **added** that the rules still **needed** to be **enforced**. The commission **said** it **suspected** criminal offences may have **been committed**, and the person responsible, Leave.EU CEO Liz Bilney, **had been referred** to the police. Related Coverage. "REMOANER SWAMP". Arron Banks, the founder of Leave.EU who **was pictured** with Donald Trump and leading Brexiteer Nigel Farage outside a gilded elevator soon after the 2016 U.S.

Construction	Count	Weight
<span style="color: green;">●</span> Verbs > Verb forms > Irre...	37	(1)
<span style="color: blue;">●</span> Verbs > Verb forms > Regu...	35	(1)

ALL CONSTRUCTIONS

🔍
↑





## Performance evaluation

Manual annotation of 351 sentences:

Linguistic target	Prec.	Rec.	F <sub>1</sub>
Yes/no questions	1.00	1.00	1.00
Irregular verbs	1.00	0.96	0.98
<i>used to</i>	0.83	1.00	0.91
Phrasal verbs	1.00	0.61	0.76
Tenses (Present Simple, ...)	0.95	0.84	0.88
Conditionals (real, unreal)	0.65	0.83	0.73
<b>Mean</b> (81 targets)	0.94	0.90	0.91
<b>Median</b> (81 targets)	1.00	0.97	0.95



## Use Cases of FLAIR

- It supports language teachers and learners in **searching** for linguistically rich reading material.
- It can serve as a **front-end** for input enhancement (Meurers et al., 2010) and exercise generation systems (Burstein et al., 2012).
- It provides a platform for Second Language Acquisition **studies** on input enrichment and enhancement.



## Use Cases of FLAIR

- It supports language teachers and learners in searching for linguistically rich reading material.
- It can serve as a front-end for input enhancement (Meurers et al., 2010) and exercise generation systems (Burstein et al., 2012).
- **It provides a platform for Second Language Acquisition studies on input enrichment and enhancement.**



---

## Studies on Input Enrichment

Studies on input enrichment with *language learners* have yielded mixed results (Trahey and White, 1993; Reinders and Ellis, 2009).

What about *language teachers*? Do they need automatic input enrichment when selecting reading material for their students?



---

## Online Study with English Teachers

- **Read** pairs of news articles (one from Bing, one from FLAIR)
- **Rate** each of the news articles
- **Select** one news article in each pair as a reading assignment



## Topic 1/10

**Topic:** Rogers Cup 2017

**Linguistic forms:** regular verbs (*typed*)

irregular verbs (*bought*)

**Start**



**Text 1**

"Pliskova reaches Rogers Cup quarters after Osaka retires ..." *(click the title to read)*

1. How relevant is the article to the topic?

(irrelevant)  1  2  3  4  5 (relevant)

2. How rich is the representation of the two target linguistic forms in the article?

(poor)  1  2  3  4  5 (good)

Next



**Text 2**

"Canada"s Bouchard ousted in first round of home tournament ..." *(click the title to read)*

1. How relevant is the article to the topic?

(irrelevant)  1  2  3  4  5 (relevant)

2. How rich is the representation of the two target linguistic forms in the article?

(poor)  1  2  3  4  5 (good)

Next





**Question**

Which news article would you give as a reading assignment to your students?

Text 1: "Pliskova reaches Rogers Cup quarters after Osaka retires ..."

Text 2: "Canada"s Bouchard ousted in first round of home tournament ..."

- Definitely Text 1    Likely Text 1    Doesn't matter    Likely Text 2    Definitely Text 2

**Submit**



---

## Repeated-measures Design

12 participants produced 240 responses by rating 20 articles in pairs:

- one top news article from Bing (original)
- one top news article from FLAIR (re-ranked)

Each pair had the same **topic**, the same **pair of linguistic forms**, was of comparable length and readability level.



---

## Topics

60 news articles from Reuters on 10 popular topics:

Game of Thrones

healthcare

street artists

Roger's Cup 2017

SpaceX

electric cars

Bitcoin

Venezuela coup

Brexit

opioid epidemic



## Pairs of Linguistic Forms

		Mean relative frequencies	
		Bing	FLAIR
frequent (95%)	regular verbs	0.012	0.020
	irregular verbs	0.012	0.019
mixed (50%)	present simple	0.011	0.014
	present continuous	0.001	0.005
infrequent (4%)	comparative d. of short adj. and adv.	0.001	0.003
	comparative d. of long adj. and adv.	0	0.001



## Sanity check

**Linguistic representation:** was FLAIR rated higher than Bing?

	Bing	FLAIR
<b>Linguistic representation</b> (1 – 5)	$M = 2.51$ $SD = 1.15$	$M = 3.22$ $SD = 1.07$

*Logistic regression* was significant:

$b = 1.89$ ,  $SE = 0.51$ ,  $p < .001$

⇒ **FLAIR > Bing** with regard to representation of linguistic forms



---

# Hypotheses

**Input enrichment:** FLAIR

**Baseline:** Microsoft Bing

**H1:** When selecting a *reading assignment*: FLAIR > Bing

**H2:** Relevance of the *content* to the topic: FLAIR < Bing

**H3:** The *more infrequent* the target linguistic forms are, the more  
FLAIR > Bing



# H1: Selecting a reading assignment: FLAIR > Bing?

Definitely Text 1    Likely Text 1    Doesn't matter    Likely Text 2    Definitely Text 2

'Doesn't matter' items (25%) were excluded from the analysis.

- English teachers selected FLAIR 71% of the time, significantly more than Bing:

$$\chi^2(1) = 16.04, p < .001$$

- 'Definitely' was selected 3 times more for FLAIR than for Bing:

$$\chi^2(1) = 12.60, p < .001$$

=> **FLAIR > Bing** when selecting a reading assignment

**A strong argument in support of automatic input enrichment.**



## H2: Relevance of content to the topic: FLAIR < Bing?

	Bing	FLAIR
<b>Relevance of content</b> (1 – 5)	$M = 3.58$ $SD = 1.00$	$M = 3.67$ $SD = 1.08$

*Logistic regression* was not significant:  
 $b = 0.53$ ,  $SE = 0.74$ ,  $p = .470$

Was it due to chance or is FLAIR = Bing with regard to content?





## H2: Relevance of content to the topic: FLAIR < Bing?

*Equivalence tests* ( $d = 0.5, \alpha = .05$ ) were significant:

$$t_1 = 4.55, p_1 < .001$$

$$t_2 = -3.19, p_2 < .001$$

$$CI [-0.13; 0.31]$$

=> **FLAIR = Bing** with regard to relevance of content to topic

**No trade-off** between content and linguistic representation in the top 20 results for popular topics in English.



## H2: Relevance of content to the topic: FLAIR < Bing?

*Equivalence tests* ( $d = 0.5, \alpha = .05$ ) were significant:

$$t_1 = 4.55, p_1 < .001$$

$$t_2 = -3.19, p_2 < .001$$

$$CI [-0.13; 0.31]$$

=> **FLAIR = Bing** with regard to relevance of content to topic

**No trade-off** between content and linguistic representation in the top 20 results for popular topics in English.



## H2: Relevance of content to the topic: FLAIR < Bing?

*Equivalence tests* ( $d = 0.5, \alpha = .05$ ) were significant:

$$t_1 = 4.55, p_1 < .001$$

$$t_2 = -3.19, p_2 < .001$$

$$CI [-0.13; 0.31]$$

=> **FLAIR = Bing** with regard to relevance of content to topic

**No trade-off** between content and linguistic representation in the top 20 results for popular topics in English.

**BUT:** Correlation of the answers to the two questions (form and content) was 0.2 ( $p < .001$ ).



## H3: The More Infrequent, the More FLAIR > Bing?

Mean preference for FLAIR (1 – 5):

Frequent	Mixed	Infrequent
Regular and irregular verbs	Present Simple and Present Continuous	Comparative degree of short and long adj. and adv.
$M = 3.46$	$M = 3.92$	$M = 3.69$
$SD = 1.39$	$SD = 1.99$	$SD = 1.30$

=> Automatic input enrichment seems particularly beneficial for targeting linguistic forms of **lower frequency levels**.



### H3: The More Infrequent, the More FLAIR > Bing?

Mean preference for FLAIR (1 – 5):

Frequent	Mixed	Infrequent
Regular and irregular verbs	Present Simple and Present Continuous	Comparative degree of short and long adj. and adv.
$M = 3.46$	$M = 3.92$	$M = 3.69$
$SD = 1.39$	$SD = 1.99$	$SD = 1.30$

=> Automatic input enrichment seems particularly beneficial for targeting linguistic forms of **lower frequency levels**.

2-way *rANOVA* was not significant:  $F(2, 90) = 0.87, p = .419$

=> **No statistically significant linear relationship** between frequency of linguistic forms and preference for FLAIR



### H3: The More Infrequent, the More FLAIR > Bing?

Check for pairwise differences (*paired 2-samples Wilcoxon tests*):

Pairs of linguistic forms	Z	p
Frequent – Mixed	157	.643
Infrequent – Mixed	128	.352
Inrequent – Frequent	217	.727

=> **No statistically significant differences** in preference for FLAIR when targeting linguistic forms of different frequencies



## Conclusion

- Teachers **preferred FLAIR** over Bing when selecting a reading assignment.
- There was **no trade-off** between relevance of content and linguistic representation: FLAIR was rated as high as Bing with regard to content
- The preference for FLAIR **did not significantly differ** when targeting frequent and infrequent linguistic forms



## Outlook

Possible empirical studies:

- A set of **independent studies** with English teachers looking at content, linguistic forms and selecting reading material separately
- A randomized controlled field study to compare the **learning outcomes of students** using web search vs. automatic input enrichment

Possible system extensions:

- Providing a **variety of contexts** in which linguistic forms are used, with their different meanings
- Integration of a component that **automatically generates exercises** targeting the selected linguistic forms





---

## Take-Home Messages

1. Develop systems for English teachers – they appreciate it!
2. Let them give you feedback – you will appreciate it!



## References

1. Burstein, J., Shore, J., Sabatini, J., Moulder, B., Holtzman, S., & Pedersen, T. (2012). The language museum system: Linguistically focused instructional authoring. ETS Research Report Series, 2012(2).
2. Chinkina, M., & Meurers, D. (2016). Linguistically aware information retrieval: Providing input enrichment for second language learners. In Proceedings of the 11th Workshop on Innovative Use of NLP for Building Educational Applications (pp. 188-198).
3. Krashen, S. (1977). The monitor model for adult second language performance. Viewpoints on English as a second language, 152-161.



## References

4. Meurers, D., Ziai, R., Amaral, L., Boyd, A., Dimitrov, A., Metcalf, V., & Ott, N. (2010, June). Enhancing authentic web pages for language learners. In Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Applications (pp. 10-18). Association for Computational Linguistics.
5. Reinders, H., & Ellis, R. (2009). The effects of two types of input on intake and the acquisition of implicit and explicit knowledge. *Implicit and explicit knowledge in second language learning, testing and teaching*, 281-302.
6. Slobin, D. I. (1985). *The crosslinguistic study of language acquisition: Vol. 2. Theoretical issues*. LEA New Jersey.
7. Trahey, M., & White, L. (1993). Positive evidence and preemption in the second language classroom. *Studies in second language acquisition*, 15(2), 181-204.



# Thank you.

Contact:

**Maria Chinkina**

[maria.chinkina@uni-tuebingen.de](mailto:maria.chinkina@uni-tuebingen.de)