



DISCOVERING PARAPHRASES FROM MOVIE SUBTITLES

Mathias Creutz

University Lecturer in Language Technology

Department of Modern Languages

Faculty of Arts

University of Helsinki

18 Oct 2017



THE DATA



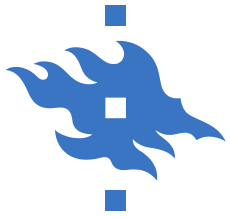
<https://www.opensubtitles.org>

<http://opus.lingfil.uu.se/>



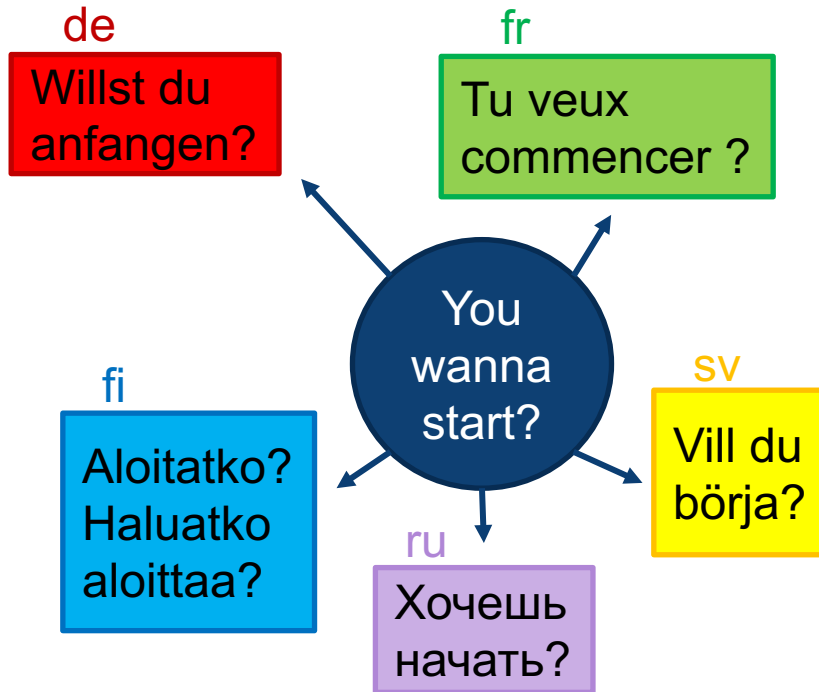
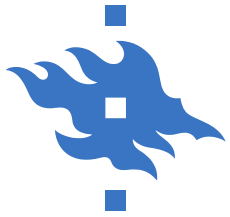
Prof. Jörg Tiedemann

ORPUS ... the open parallel corpus



EXAMPLE: FINDING PARAPHRASES FOR “YOU WANNA START?”







Do you want to go first?
Or else you can start.
You wanna break?

Willst du
anfangen?

Tu veux
commencer ?

Do you want to go first?
Do you want to proceed?
No, I think you should.
Okay, you wanna go first?
Wh-When did you talk to him?
You wanna start us off?
You want to be first?
You want to go first?

You
wanna
start?

Do you wanna start?
Do you wanna volley for serve?
Do you want to go first?
Do you want to start?
Hey, you want to tap out?
The new dad's gonna.
Want to go first?
You wanna get started?
You wanna go first?
You want to get into this right now?

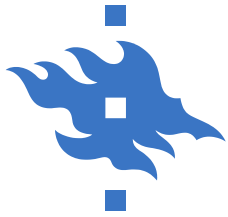
Aloitatko?
Haluatko
aloittaa?

Vill du
börja?

D'you want to go first?
Do you know the opening?
Do you want to go first?
Do you want to start?
This calls for a joint.
Why don't you start us off.
Would you like to start?
You wanna go first?
You want to go first?

Хочешь
начать?

Did you wanna go first?
Do you want to go first?
Want to start?



Do you want to go first?
 Or else you can start.
 You wanna break?

Willst du anfangen?

Tu veux commencer ?

Do you want to go first?
 Do you want to proceed?
 No, I think you should.
 Okay, you wanna go first?
 Wh-When did you talk to him?
 You wanna start us off?
 You want to be first?
You want to go first?

Do you wanna start?
 Do you wanna volley for serve?
Do you want to go first?
Do you want to start?
 Hey, you want to tap out?
 The new dad's gonna.
 Want to go first?
 You wanna get started?
You wanna go first?
 You want to get into this right now?

Aloitatko? Haluatko aloittaa?

You wanna start?

Vill du börja?

Хочешь начать?

D'you want to go first?
 Do you know the opening?
Do you want to go first?
Do you want to start?
 This calls for a joint.
 Why don't you start us off.
 Would you like to start?
You wanna go first?
You want to go first?

Did you wanna go first?
Do you want to go first?
 Want to start?



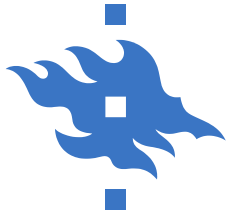
INTERESTING OR NOT?



Who's Elise? ↔ Who is Elise?

Good job, buddy! ↔ Nice work, man!

I have goose flesh. ↔
The hair's standing up on my arms.



THANK YOU!