

Simple, Lexicalized Choice of Translation Timing for Simultaneous Speech Translation

Tomoki Fujita, Graham Neubig, Sakriani Sakti,
Tomoki Toda, Satoshi Nakamura

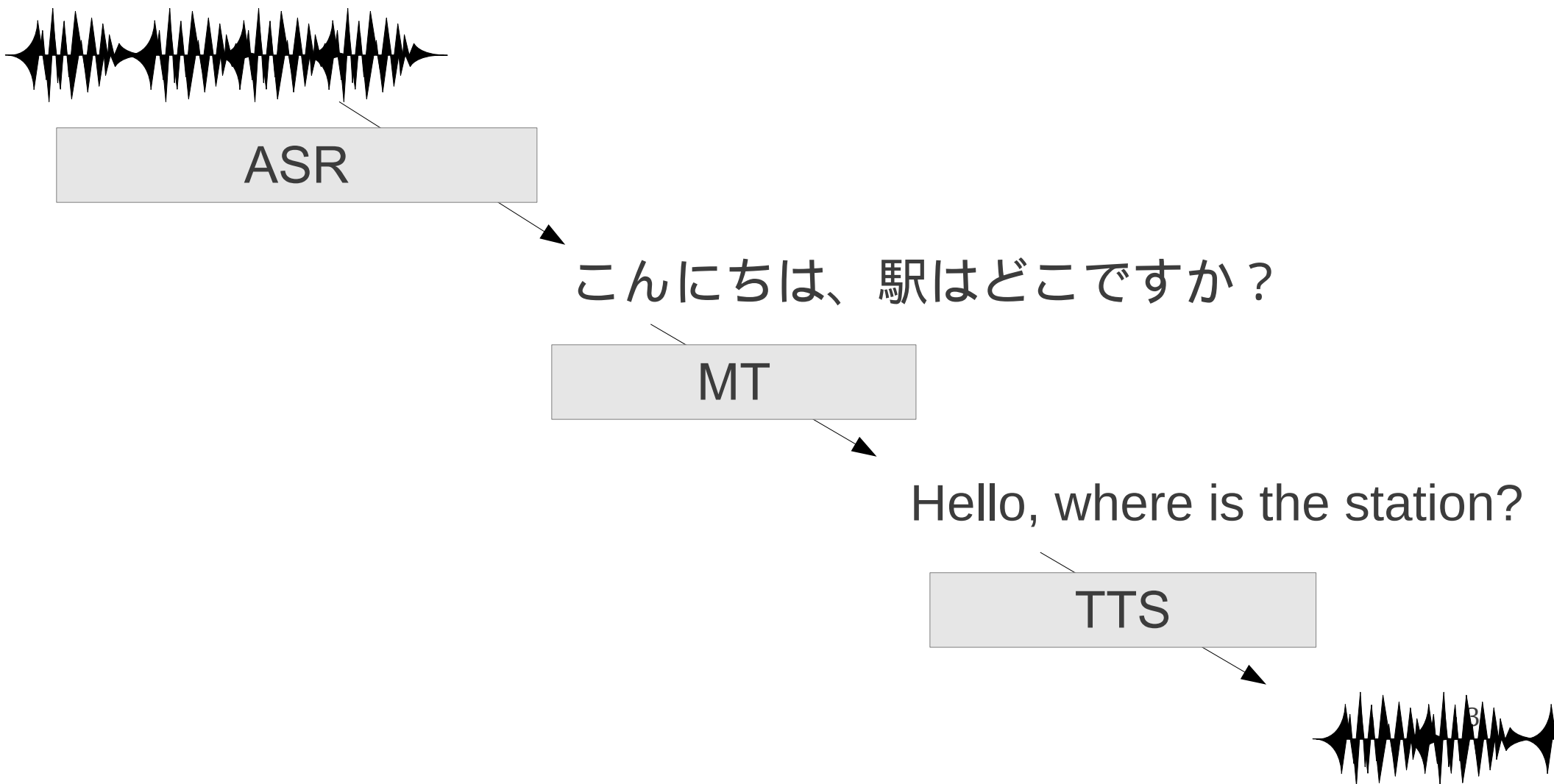
Graduate School of Information Science,
Nara Institute of Science and Technology



Background

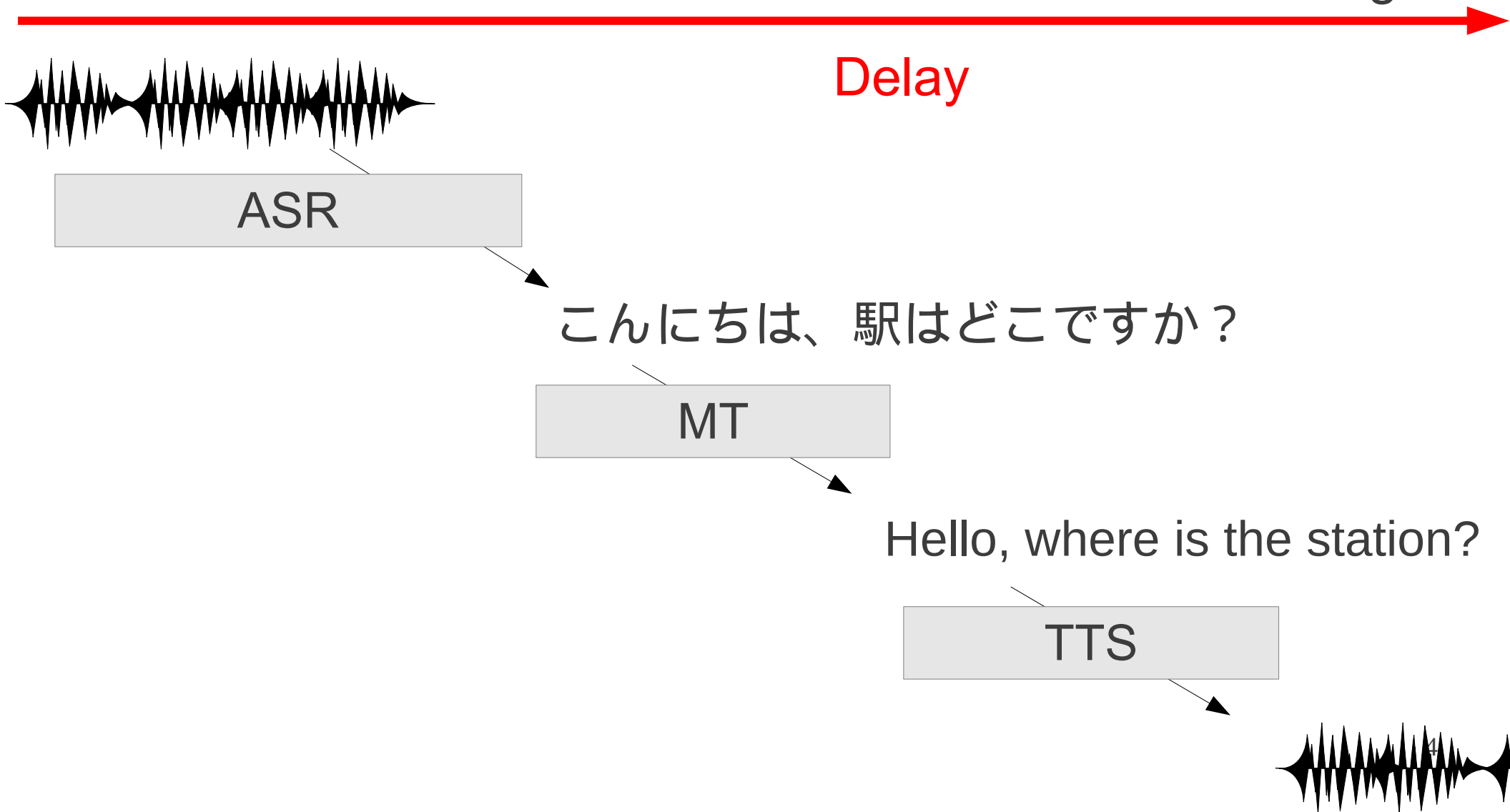
Speech Translation Systems

- Translate speech from source language to target



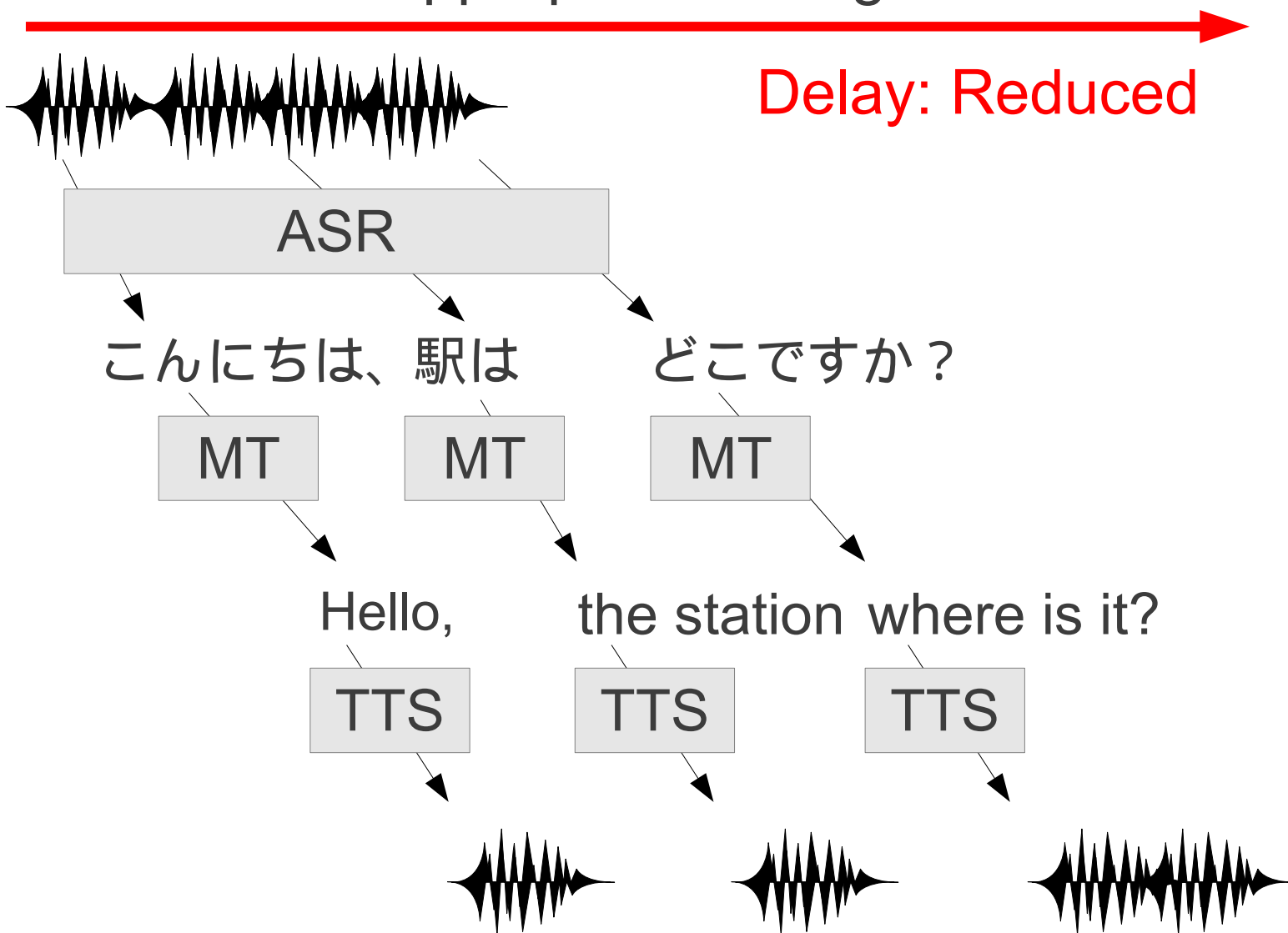
Problem: Delay

- Wait for the whole utterance to end before translating



Solution: Divide into Smaller Chunks

- Choose appropriate timing to start translation



Previous Work: Incremental Dependency Parsing/Manual Rules [Ryu+ 04]

- Utilize knowledge of English/Japanese to derive rules

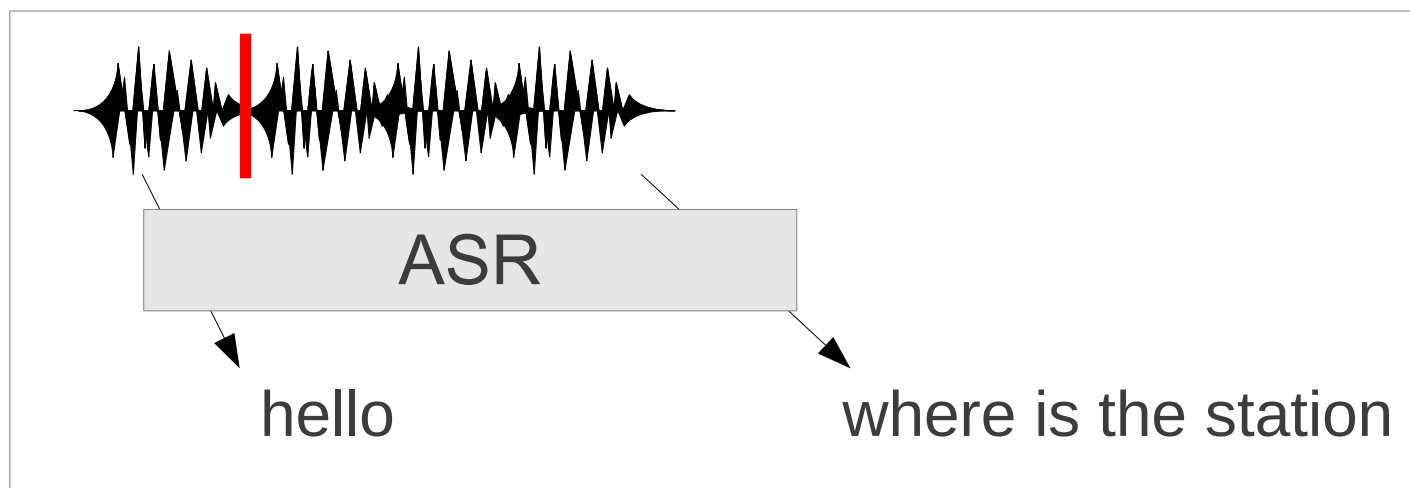


- Requires a bilingual linguist to design rules
- Requires an accurate incremental dependency parser

Previous Work: Division on Pauses

[Fugen+ 08, Bangalore+ 12]

- Simply divide on short pauses in the utterance



- - Cannot capture relationship between languages
- - Result will greatly change with speech speed, disfluencies

Proposed Method

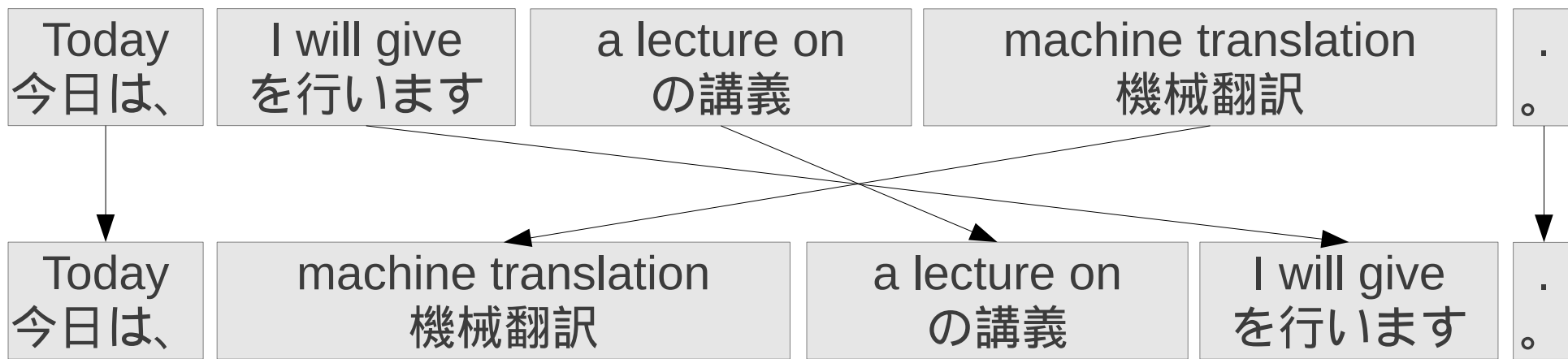
- Utilize the TM directly to choose translation timing
 - + Can be constructed automatically
 - + Uses information about the language pair
 - + Very simple to implement
- Specifically:
 - Choose translation timing at the end of each **phrase** in the phrase table
 - Utilize **reordering probabilities** to adjust granularity
 - **Adapt** the language model to the translation task

Preliminaries

Phrase Based Machine Translation

- Divide the sentence into small phrases and translate

Today I will give a lecture on machine translation .



今日は、機械翻訳の講義を行います。

- Score translations with **translation model (TM)**, **reordering model (RM)**, and **language model (LM)**

Translation Model Creation

- Perform automatic alignment of bitext
- From aligned text, extract phrases for translation

ホ
テ 受
ルの付

the
hotel
front
desk

ホテルの → hotel
ホテルの → the hotel
受付 → front desk
ホテルの受付 → hotel front desk
ホテルの受付 → the hotel front desk

Lexicalized Reordering Model

- Probabilistically models reorderings for increased accuracy of translation
- Given **current phrase** and **next phrase**:

Monotone:

背 の 高い 男
 the tall man

Swap:

太郎 を 訪問 した
 visited Taro

Discontinuous Right:

私 は 太郎 を 訪問した
 I visited Taro

Discontinuous Left:

背 の 高い 男 を 訪問した
 visited the tall man

- “monotone” + “discontinuous right” = “right probability”

Proposed Method

Method One:

Choosing Translation Timing with Phrases

- Input words one at a time from ASR
- While words exist in phrase table, don't translate yet

Method One: Choosing Translation Timing with Phrases

- Input words one at a time from ASR
- While words exist in phrase table, don't translate yet

Phrase Table

hello → こんにちは

where → どこ

where is → どこですか

the → その

the station → 駅

Input String

hello

where

is

the

station

Method One: Choosing Translation Timing with Phrases

- Input words one at a time from ASR
- While words exist in phrase table, don't translate yet

Phrase Table

hello → こんにちは

where → どこ

where is → どこですか

the → その

the station → 駅

Input String

hello

where

is

the

station

Method One: Choosing Translation Timing with Phrases

- Input words one at a time from ASR
- While words exist in phrase table, don't translate yet

Phrase Table

hello → こんにちは

where → どこ

where is → どこですか

the → その

the station → 駅

Input String

hello

where

is

the

station

“hello”
phrase exists



wait

Method One: Choosing Translation Timing with Phrases

- Input words one at a time from ASR
- While words exist in phrase table, don't translate yet

Phrase Table

hello → こんにちは

where → どこ

where is → どこですか

the → その

the station → 駅

Input String

hello

where

is

the

station

“hello”
phrase exists



wait

Method One: Choosing Translation Timing with Phrases

- Input words one at a time from ASR
- While words exist in phrase table, don't translate yet

Phrase Table

hello → こんにちは

where is → どこですか

the station → 駅

where → どこ

the → その

Input String

hello

where

is

the

station

“hello”
phrase exists

“hello where”
phrase missing

↓
wait

↓
translate
“hello”

Method One: Choosing Translation Timing with Phrases

- Input words one at a time from ASR
- While words exist in phrase table, don't translate yet

Phrase Table

hello → こんにちは

where → どこ

where is → どこですか

the → その

the station → 駅

Input String

hello

where

is

the

station

“hello”
phrase exists

↓
wait

“hello where”
phrase missing

↓
translate
“hello”

Method One: Choosing Translation Timing with Phrases

- Input words one at a time from ASR
- While words exist in phrase table, don't translate yet

Phrase Table

hello → こんにちは

where is → どこですか

the station → 駅

where → どこ

the → その

Input String

hello

where

is

the

station

“hello”
phrase exists

↓
wait

“hello where”
phrase missing

↓
translate
“hello”

“where is”
phrase exists

↓
wait

Method One: Choosing Translation Timing with Phrases

- Input words one at a time from ASR
- While words exist in phrase table, don't translate yet

Phrase Table

hello → こんにちは

where → どこ

where is → どこですか

the → その

the station → 駅

Input String

hello

where

is

the

station

“hello”
phrase exists

↓
wait

“hello where”
phrase missing

↓
translate
“hello”

“where is”
phrase exists

↓
wait

Method One: Choosing Translation Timing with Phrases

- Input words one at a time from ASR
- While words exist in phrase table, don't translate yet

Phrase Table

hello → こんにちは where is → どこですか the station → 駅
 where → どこ the → その

Input String

hello where ————— is the station

“hello”
phrase exists

↓
wait

“hello where”
phrase missing

↓
translate
“hello”

“where is”
phrase exists

↓
wait

“where is the”
phrase missing

↓
translate
“where is”

Method One: Choosing Translation Timing with Phrases

- Input words one at a time from ASR
- While words exist in phrase table, don't translate yet

Phrase Table

hello → こんにちは

where is → どこですか

the station → 駅

where → どこ

the → その

Input String

hello

where ————— is

the

station

“hello”
phrase exists

“hello where”
phrase missing

“where is”
phrase exists

“where is the”
phrase missing

↓
wait

↓
translate
“hello”

↓
wait

↓
translate
“where is”

Method One: Choosing Translation Timing with Phrases

- Input words one at a time from ASR
- While words exist in phrase table, don't translate yet

Phrase Table

hello → こんにちは

where is → どこですか

the station → 駅

where → どこ

the → その

Input String

hello

where ——— is

the ——— station

“hello”
phrase exists

“hello where”
phrase missing

“where is”
phrase exists

“where is the”
phrase missing

“the station”
utterance ends

↓
wait

↓
translate
“hello”

↓
wait

↓
translate
“where is”

↓
translate
“the station”

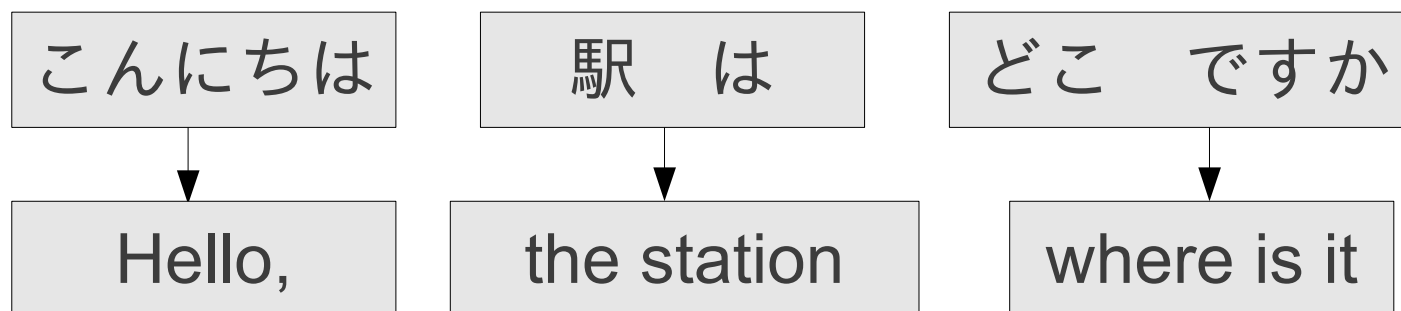
Problem with Method One

- Has the potential to degrade translation accuracy:

Normal phrase-based translation:



Translation with early timing:

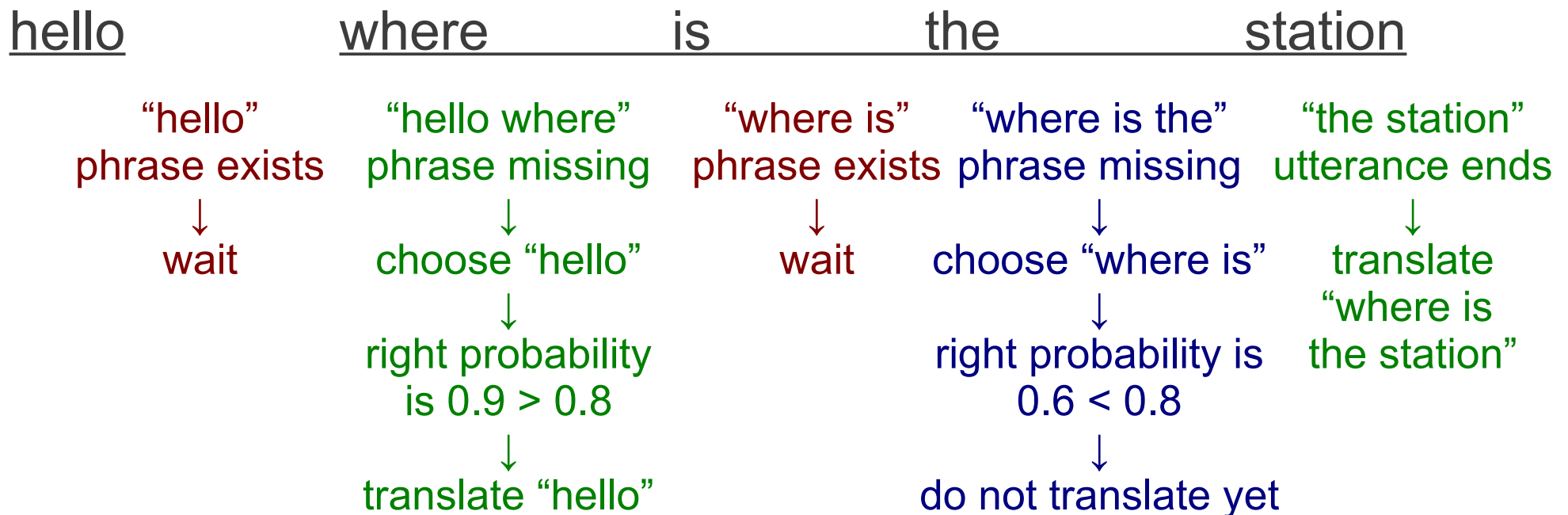


Method Two:

Adjusting Timing with Reordering Probabilities

- First, temporarily choose strings according to method one
- Next, if that phrase's **right probability** exceeds a threshold, actually translate the words in the cache

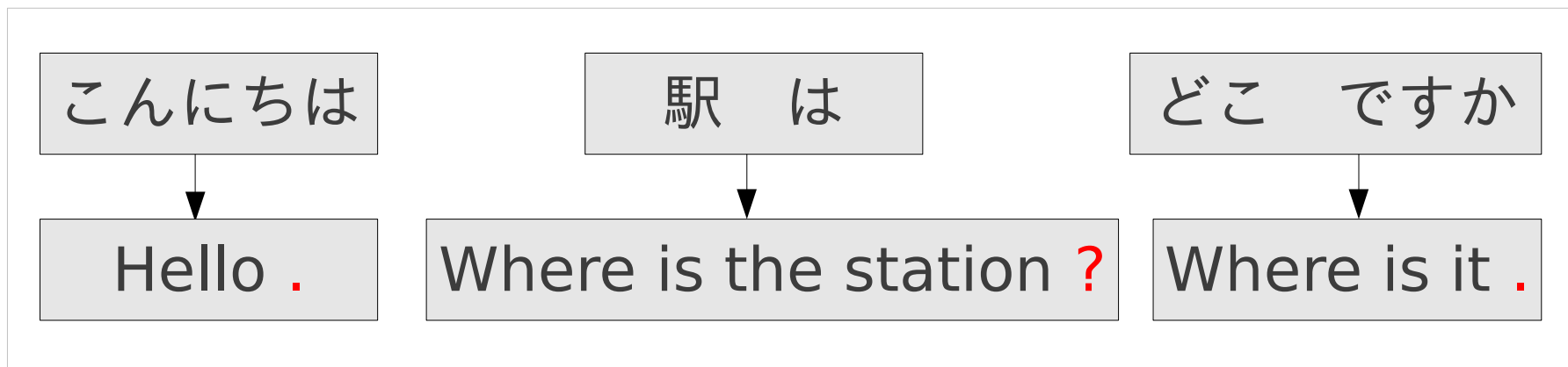
Example (threshold = 0.8):



- Threshold 1.0 = traditional, 0.0 = method one

Problem with Method Two

- LMs are traditionally trained on sentences
 - This is not appropriate for translating shorter chunks
 - e.g.: The translator will try to “finish” sentences

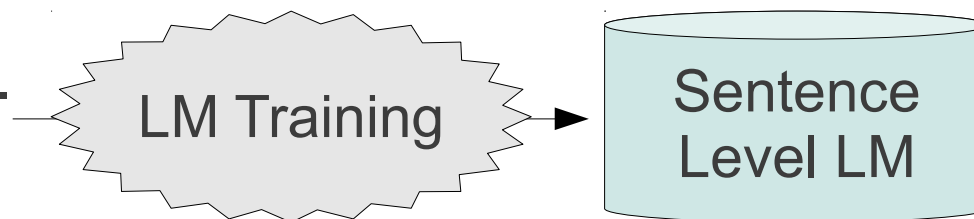


Method Three: Language Model Adaptation

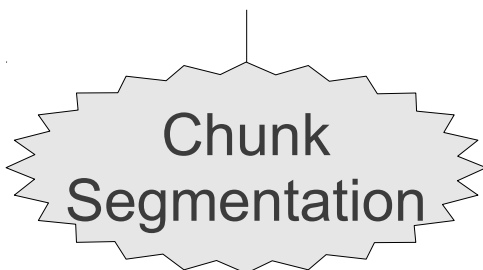
- Before learning the language model, **split the training data** according to the same criterion

Traditional LM Training

Hello, where is the station.
My name is John.



Proposed Method



Hello,
where is
the station.
My name
is John.

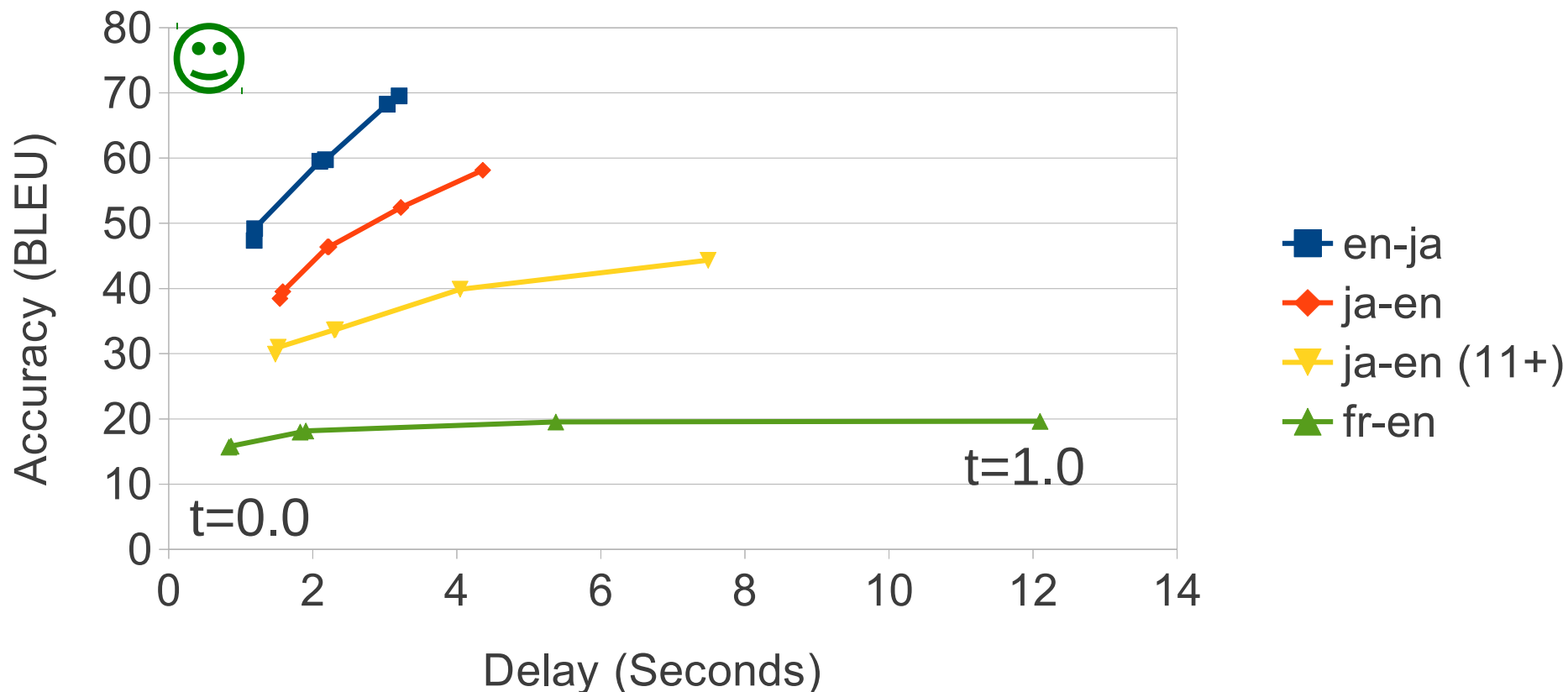


Experiments

Experimental Setup

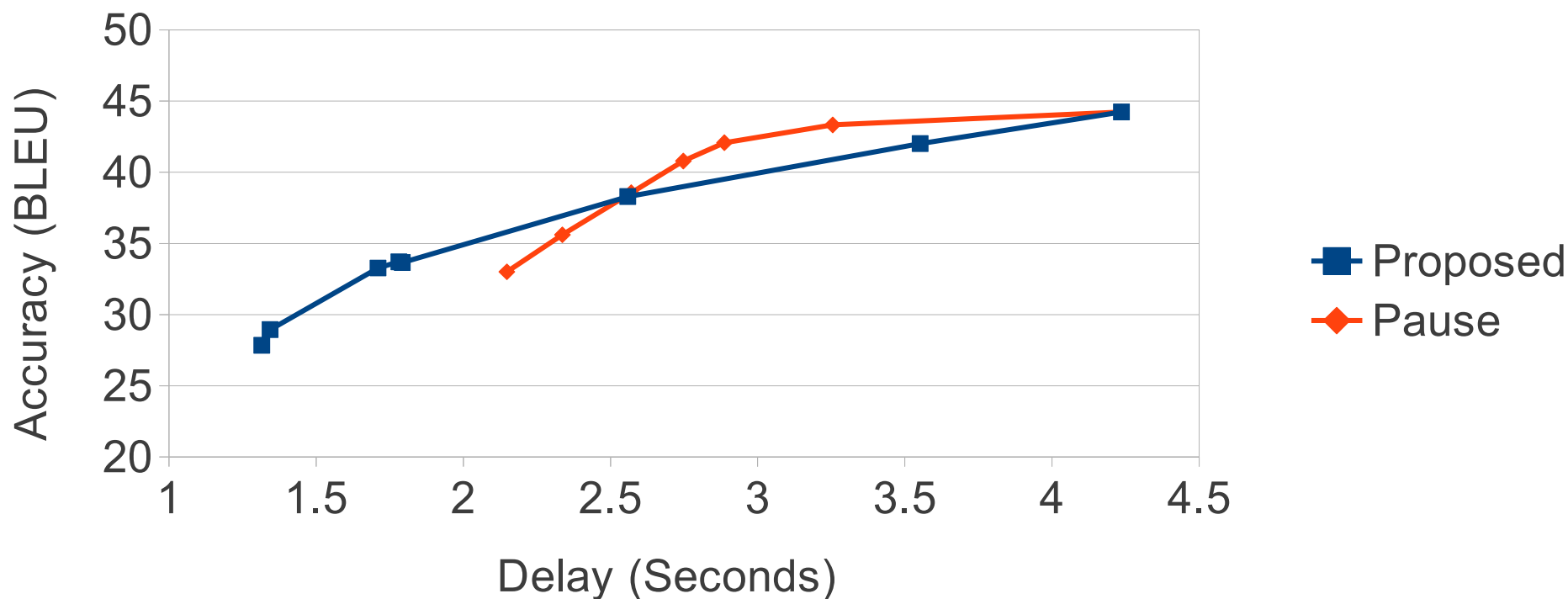
- **Four Types of Experiments:**
 - Japanese-English BTEC Travel Conversation (ja-en)
 - Japanese-English BTEC with 11+ Words (ja-en 11+)
 - English-Japanese BTEC Travel Conversation (en-ja)
 - French-English WMT News (fr-en)
- **Evaluation Measures:**
 - Accuracy
 - 14-ref BLEU for BTEC, 1-ref BLEU for News
 - Manually-graded acceptability
 - Delay (Seconds)

Result One: Comparison Across Settings



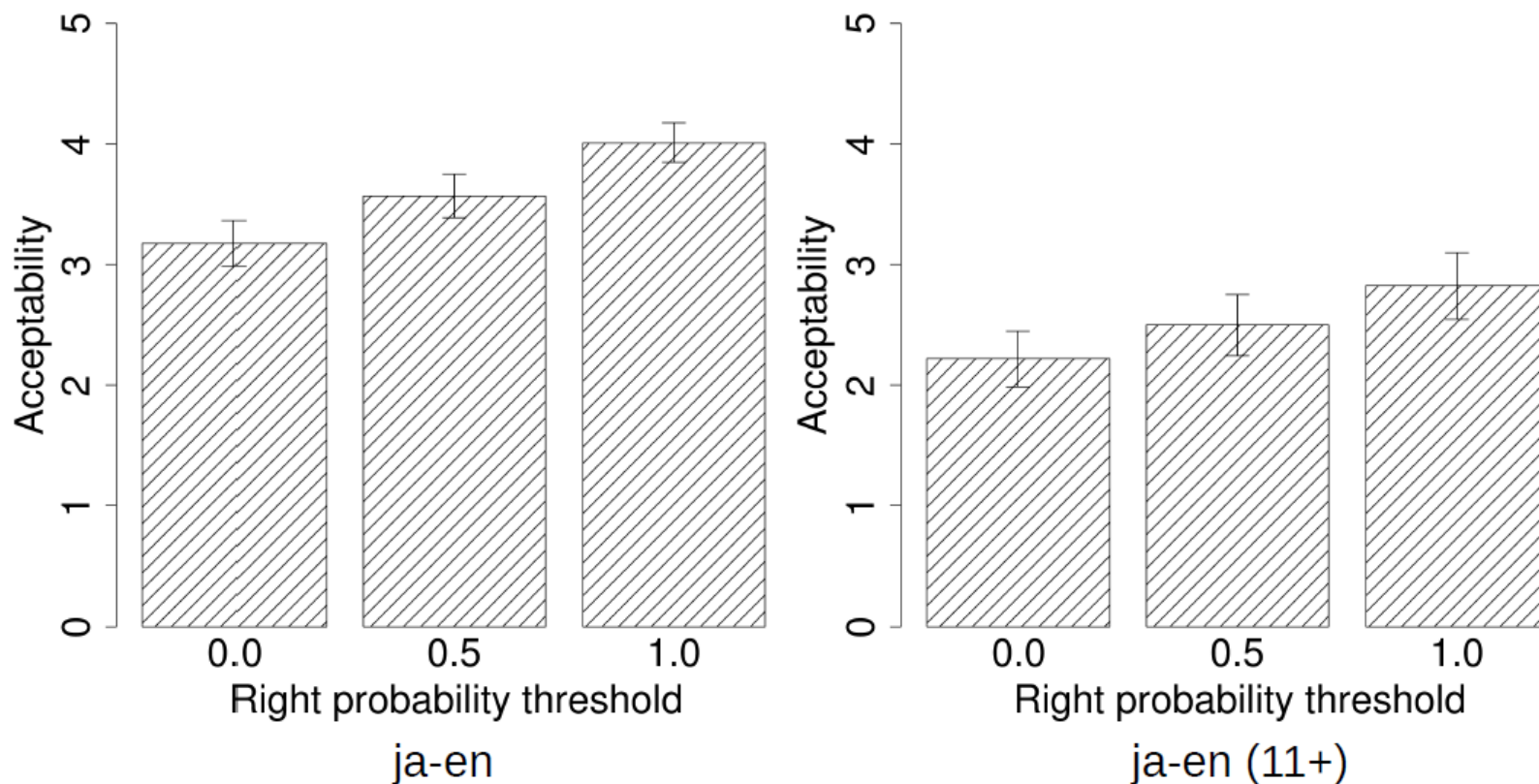
- Delay decreases in all settings
- Better delay/accuracy tradeoff for long sentences, similar languages

Result Two: Compare with Pause-based Segmentation



- In faster settings proposed method best
- In slower settings pause-based method best

Manual Evaluation



- Decrease in manual evaluation as well, but less obvious than evaluated by BLEU

Conclusion

Conclusion

- Proposed method for choosing timing in speech translation using phrase table and reordering model
 - **Considers reordering tendencies** across languages
 - **Simple and language independent**
 - **Competitive accuracy**
- Future work:
 - **Combination** of prosodic, reordering, and syntactic cues
 - How do we **evaluate** translations?

Thank You!

Example:

サーフィン に いい 場所 を 教え て ください

please tell me a good surfing place

サーフィン に | いい 場所 を 教え て ください

for surfing | please tell me a good place