

# Designing and Building a Japanese Controlled Language for the Automotive Domain

Toward the Development of a Writing Assistant Tool

---

---

**Takahiro Makino**, Rei Miyata, Seo Sungwon and Satoshi Sato  
Nagoya University, Nagoya, Japan

[makino.takahiro.v0@s.mail.nagoya-u.ac.jp](mailto:makino.takahiro.v0@s.mail.nagoya-u.ac.jp)

# Problem and Solution

## ■ Automotive technical documents

- written by many writers and editors

## ■ Problem

- Writing variations
- Degraded searchability and reusability
- negative impact on translation memory tools

## ■ Solution

- Controlled language
- Writing assistant tool

**delete?  
remove?**



# Writing Variations in Japanese

## Lexical variation

Futatabi tenken o okonau.

Saido tenken o okonau.

## Word location variation

Tenken o saido okonau

Saido ..... tenken o okonau

(Trans.: Perform the check **again**)

# Writing Variations in Japanese

## Lexical variation

Futatabi tenken o okonau.

word substitution

Saido tenken o okonau.

## Word location variation

Tenken o saido okonau

Saido ..... tenken o okonau

(Trans.: Perform the check **again**)

# Writing Variations in Japanese

## Lexical variation

Futatabi tenken o okonau.

word substitution

Saido tenken o okonau.

## Word location variation

Tenken o saido okonau

word reordering

Saido tenken o okonau

(Trans.: Perform the check **again**)

# Demo

## ■ Word substitution

センテンス編集

メモ あとで修正。

トラブルシュートを実施する前に、この回路のヒューズを点検する。

文字数：31

トラブルシュートを**実施する前に**、  
この回路のヒューズを点検する。

順序を整える

一括訂正

# Demo

## ■ Word substitution

センテンス編集

メモ あとで修正。

トラブルシュートを実施する前に、この回路のヒューズを点検する。

文字数：31

トラブルシュートを**実施する前に**、  
この回路のヒューズを点検する。

順序を整える

一括訂正

# Demo

## ■ Word reordering

センテンス編集

メモ

作業する前に、S R Sエアバッグシステムの注意事項を必ず確認する。

文字数：32

作業する前に、  
S R Sエアバッグシステムの注意事項を必ず  
確認する。

順序を整える

一括訂正

順序訂正後

作業する前に、必ずS R Sエアバッグシステム  
の注意事項を確認する。



# Demo

## ■ Word reordering

センテンス編集



メモ

作業する前に、S R Sエアバッグシステムの注意事項を必ず確認する。

文字数：32

作業する前に、  
S R Sエアバッグシステムの注意事項を必ず  
確認する。

順序を整える

一括訂正

順序訂正後

作業する前に、必ずS R Sエアバッグシステム  
の注意事項を確認する。

# Writing Variations in Japanese

## Lexical variation

Futatabi tenken o okonau.

word substitution

Saido tenken o okonau.

## Word location variation

Tenken o saido okonau

word reordering

Saido tenken o okonau

(Trans.: Perform the check **again**)

# Writing Variations in Japanese

## Lexical variation

*Futatabi* *tenken o okonau.*

word substitution

*Saido* *tenken o okonau.*

## Word location variation

*Tenken o* *saido* *okonau*

word reordering

*Saido* *tenken o okonau*

(Trans.: Perform the check **again**)

We designed and built a **controlled language (lexicon)** that supports **word substitution** and **word reordering**

# Overview

---

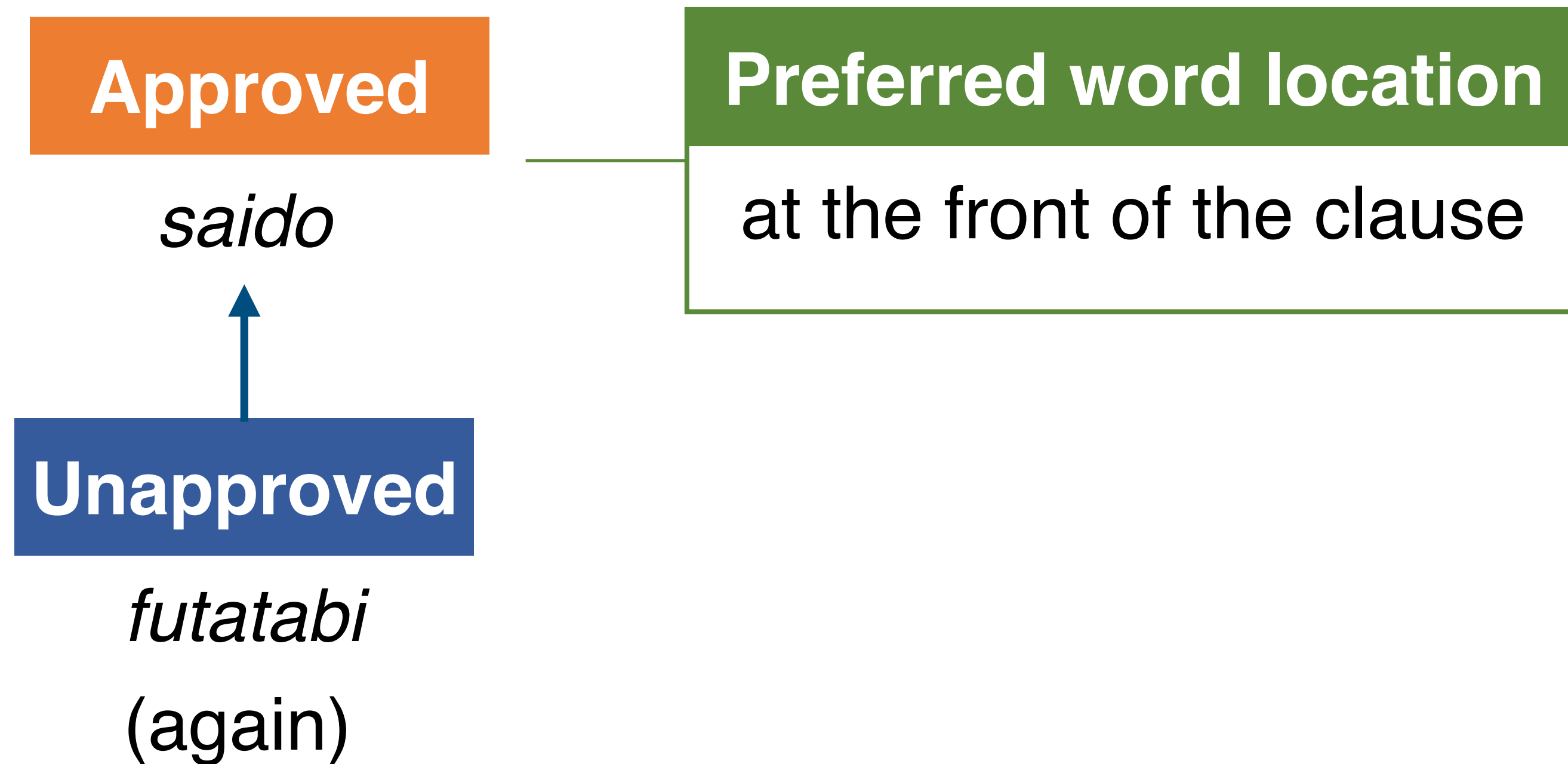
1. Design of a controlled language
2. Building procedure and results

# 1. Design

---

# Principle of Design

- For word substitution: **Approved/unapproved word list** (Warburton 2014; ASD 2021)
  - One meaning should correspond to one form
  - cf. controlled lexicon: approved word list
- For word reordering : **Preferred word location** information



# Components

- Five components that can widely cover sentence elements in the target domain

the clip A    from the vehicle    be removed    have    when    surely    with a new one    replace    be sure to

[クリップA を] [車両 から] 取りはずし た 場合は、 必ず [新品 に] 交換する こと

*kurippu A o sharyou kara torihazushi ta baai wa kanarazu shinpini koukansuru koto*

- case frame
- case element
- adverbial modifier
- sentence-ending function
- connective

Our components are based on **units of writing**

- cf. ASD-STE100 uses eight POS:  
noun, verb, adjective, adverb, pronoun,  
article, preposition, conjunction

# Components

- Five components that can widely cover sentence elements in the target domain

[the clip A] [from the vehicle] **be removed** have when surely [with a new one] **replace** be sure to

[クリップA **を**] [車両 **から**] **取りはずし** た 場合は、必ず [新品 **に**] **交換する** こと

[kurippu A **o**] [sharyou **kara**] **torihazushi** ta baai wa kanarazu [shinpin **ni**] **koukansuru** koto

- **case frame**

(Miyata/Sugino 2020)

- case element
- adverbial modifier
- sentence-ending function
- connective

verbs with canonical orders for their argument slots



# Components

- Five components that can widely cover sentence elements in the target domain

[the clip A] [from the vehicle] be removed have when surely [with a new one] replace be sure to

[クリップA を] [車両 から] 取りはずし た 場合は、必ず [新品 に] 交換する こと

[kurippu A o] [sharyou kara] torihazushi ta baai wa kanarazu [shinpin ni] koukansuru koto

- case frame
- **case element**
- adverbial modifier
- sentence-ending function
- connective

nouns or noun phrases that can fill argument slots

# Components

- Five components that can widely cover sentence elements in the target domain

the clip A    from the vehicle    be removed    have    when    **surely**    with a new one    replace    be sure to

[クリップA を] [車両 から] 取りはずし た 場合は、 **必ず** [新品 に] 交換する こと

*kurippu A o    sharyou kara    torihazushi    ta    baai wa    **kanarazu**    shinpin ni    koukansuru    koto*

- case frame
- case element
- **adverbial modifier**
- sentence-ending function
- connective

adverbs or adverbial phrases  
with their preferred locations

# Components

- Five components that can widely cover sentence elements in the target domain

the clip A    from the vehicle    be removed    **have**    when    surely    with a new one    replace    **be sure to**

[クリップA を] [車両 から] 取りはずし **た** 場合は、必ず [新品 に] 交換する **こと**

*kurippu A o    sharyou kara    torihazushi    ta    baai wa    kanarazu    shinpin ni    koukansuru    koto*

- case frame
- case element
- adverbial modifier
- **sentence-ending function**
- connective

sequences of functional words  
attached to sentence-ending main verbs

# Components

- Five components that can widely cover sentence elements in the target domain

the clip A    from the vehicle    be removed    have    **when**    surely    with a new one    replace    be sure to

[クリップA を] [車両 から] 取りはずし た **場合は、** 必ず [新品 に] 交換する こと

*kurippu A o    sharyou kara    torihazushi    ta    baai wa    kanarazu    shinpin ni    koukansuru    koto*

- case frame
- case element
- adverbial modifier
- sentence-ending function
- **connective**

conjunctions and conjunctive phrases  
that indicate inter-clause or  
inter-sentence relationships

# Components

- Five components that can widely cover sentence elements in the target domain

the clip A    from the vehicle    be removed    have    when    surely    with a new one    replace    be sure to

[クリップA を] [車両 から] 取りはずし た 場合は、 必ず [新品 に] 交換する こと

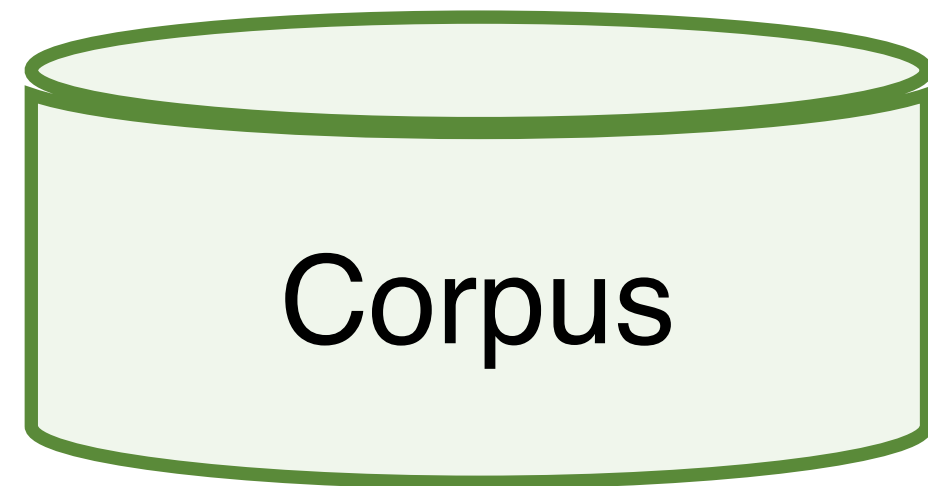
*kurippu A o    sharyou kara    torihazushi    ta    baai wa    kanarazu    shinpin ni    koukansuru    koto*

- case frame
- case element
- **adverbial modifier**
- **sentence-ending function**
- **connective**

# 2. Lexicon Building

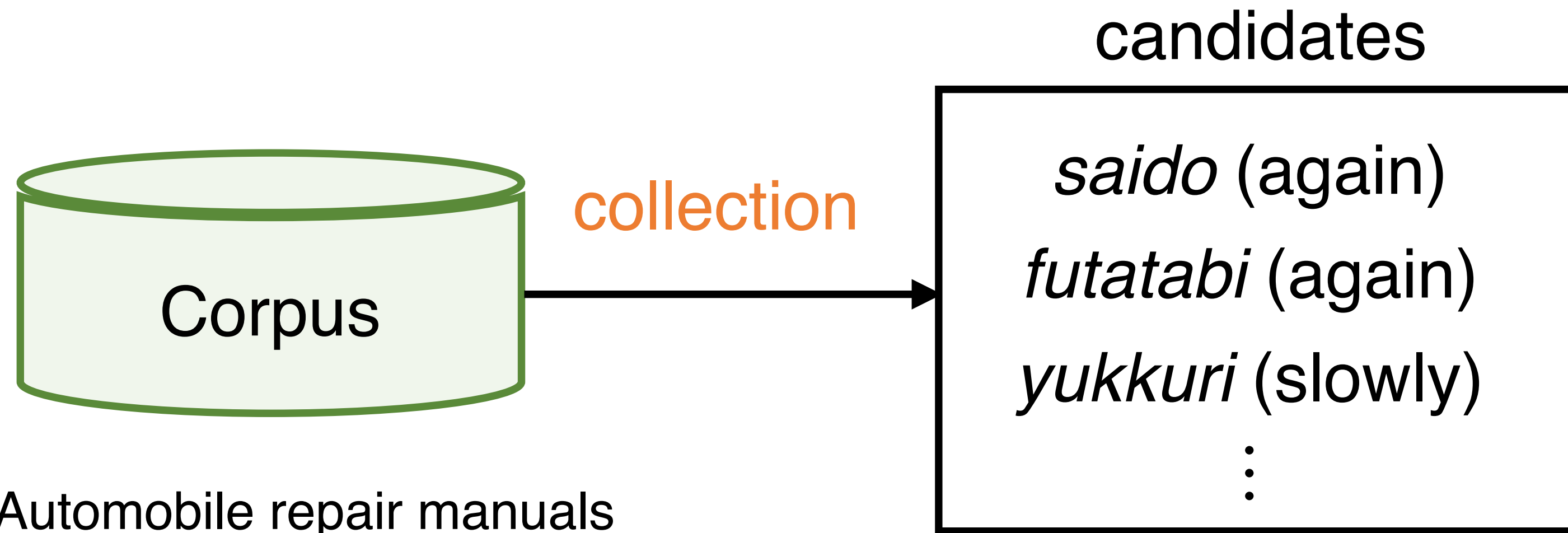
---

# General Methodology: Corpus-based



- Automobile repair manuals
- 1,053,111 sentence tokens
- 158,383 sentence types

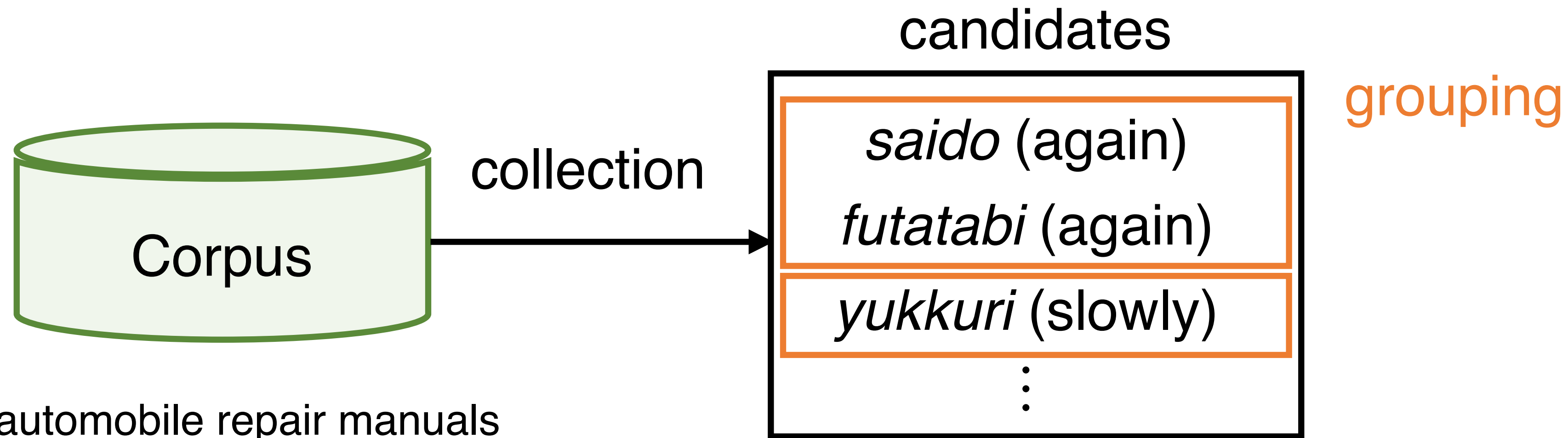
# General Methodology: Corpus-based



- Automobile repair manuals
- 1,053,111 sentence tokens
- 158,383 sentence types

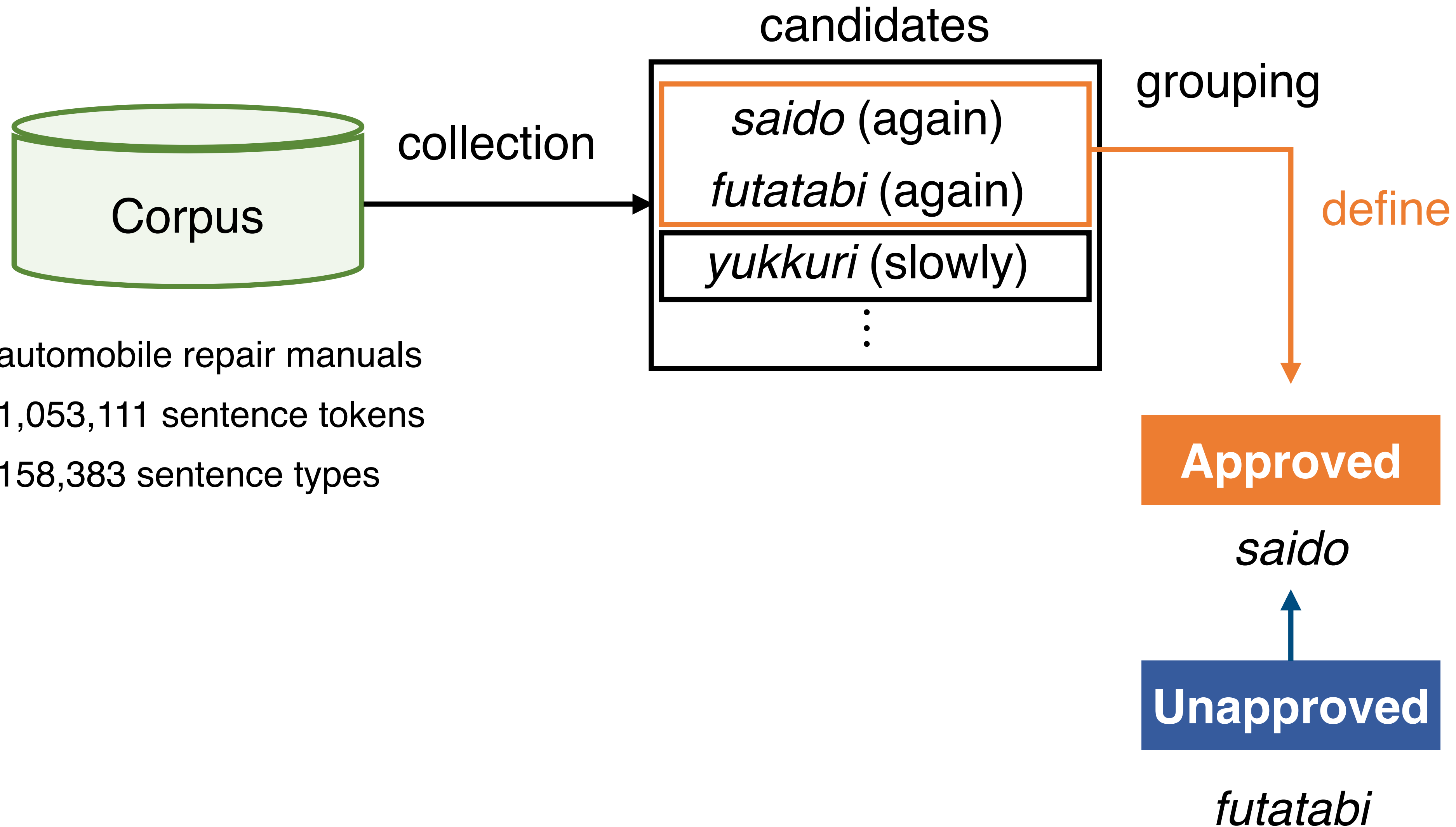


# General Methodology: Corpus-based

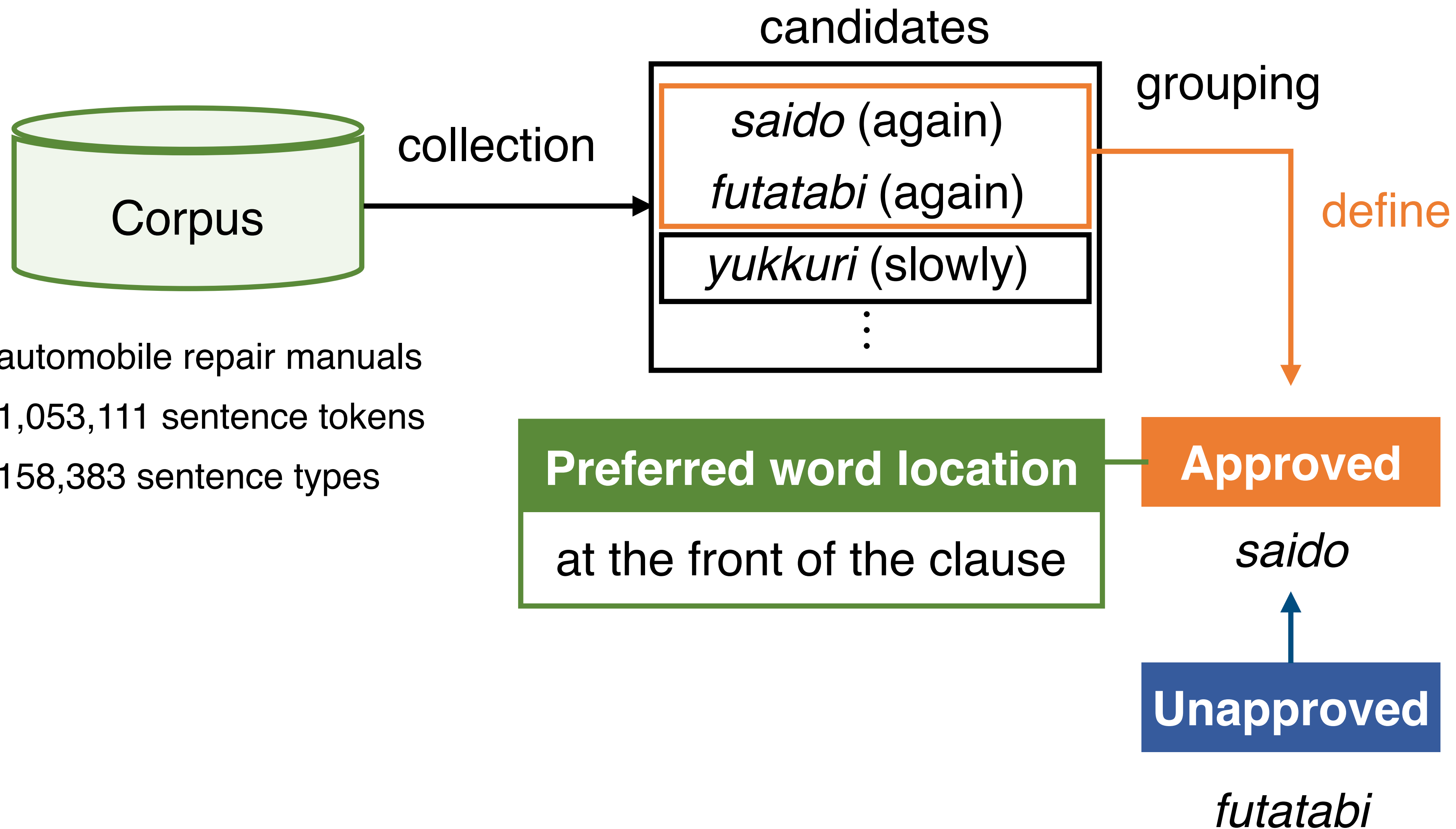


- automobile repair manuals
- 1,053,111 sentence tokens
- 158,383 sentence types

# General Methodology: Corpus-based



# General Methodology: Corpus-based



- automobile repair manuals
- 1,053,111 sentence tokens
- 158,383 sentence types

# Lexicon Building: Adverbial Modifiers

the clip A    from the vehicle    be removed    have    when    **surely**    with a new one    replace    be sure to  
[クリップA を] [車両 から] 取りはずし た 場合は、 **必ず** [新品 に] 交換する こと  
*kurippu A o    sharyou kara    torihazushi    ta    baai wa    kanarazu    shinpin ni    koukansuru    koto*

- case frame
- case element
- **adverbial modifier**
- sentence-ending function
- connective

adverbs or adverbial phrases  
with their preferred locations

# Types of Adverbial Modifiers

## ■ Fixed type

- 必ず/*kanarazu* (surely)
- 同量ずつ/*douryou zutsu* (by an equal amount)

## ■ Variable type

- 約[X]分間/*yaku [X] funkan* (for approximately [X] minutes)
- [X]後/[X] *go* (after [X])

# Extraction Method

- Heuristic rules to capture adverbial modifiers

- based on the types of **last morpheme** (if it is a particle, **the second last**)

**#1: adverb      #2: continuative form of adjective      #3: temporal noun used as an adverb**

ゆっくり

*yukkuri*

(slowly)

無理に

*murini*

(by force)

再度

*saido*

(again)

**#4: suffix to make an adverb**

定期的に

*teiki-tekini*

(periodically)

**#5: temporal suffix**

作業前に

*sagyou-maeni*

(before proceeding with work)

**#6: numeral suffix**

15秒間

*15 byoukan*

(for 15 seconds)

# Building Procedure

## Classification

extracted candidates

- *saido* (again)
  
- *yaku 15 funkan*  
(for approximately 15 minutes)
- *yaku 5 funkan*  
(for approximately 5 minutes)

# Building Procedure

## Classification

extracted candidates

**fixed**

- *saido* (again)

**variable**

- *yaku 15 funkan*  
(for approximately 15 minutes)
- *yaku 5 funkan*  
(for approximately 5 minutes)



# Building Procedure

## Classification

extracted candidates

**fixed**

- *saido* (again)

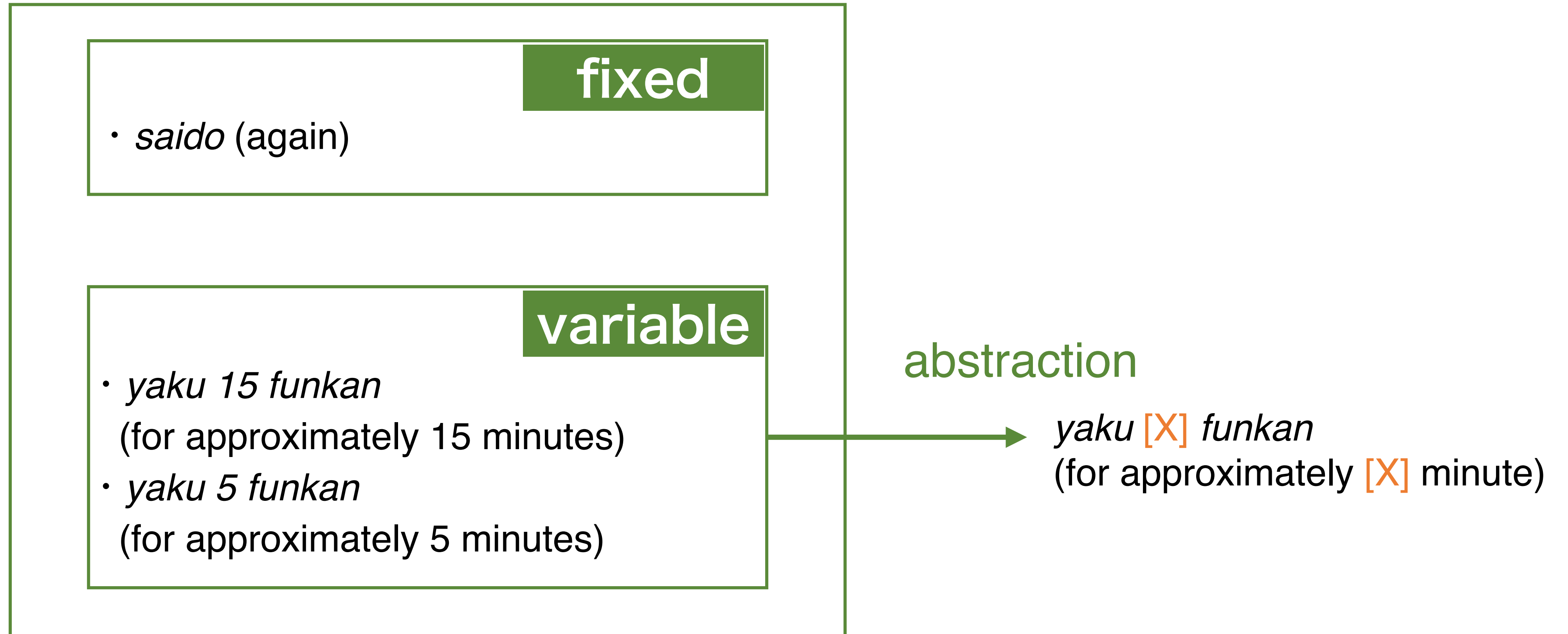
**variable**

- *yaku 15 funkan*  
(for approximately 15 minutes)
- *yaku 5 funkan*  
(for approximately 5 minutes)

# Building Procedure

## Classification

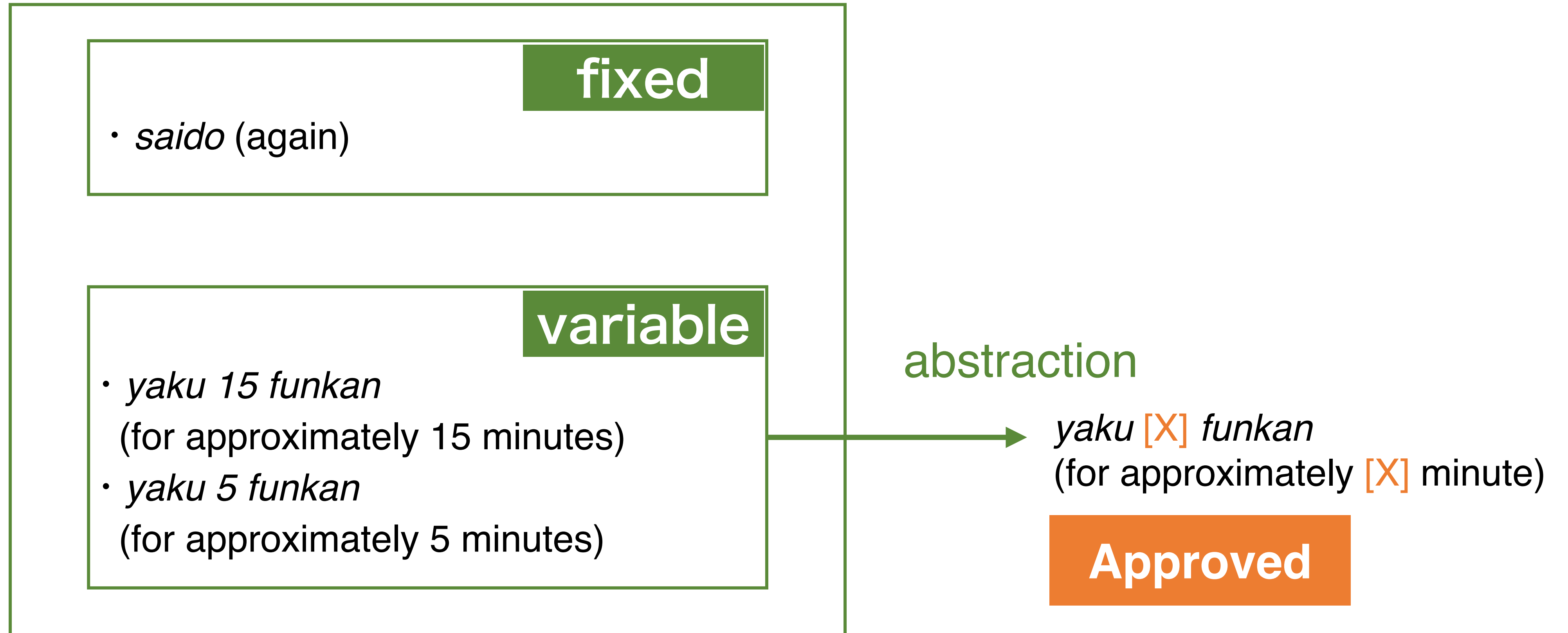
extracted candidates



# Building Procedure

## Classification

extracted candidates



# Building Procedure

---

## Definition of preferred word location

**Approved**

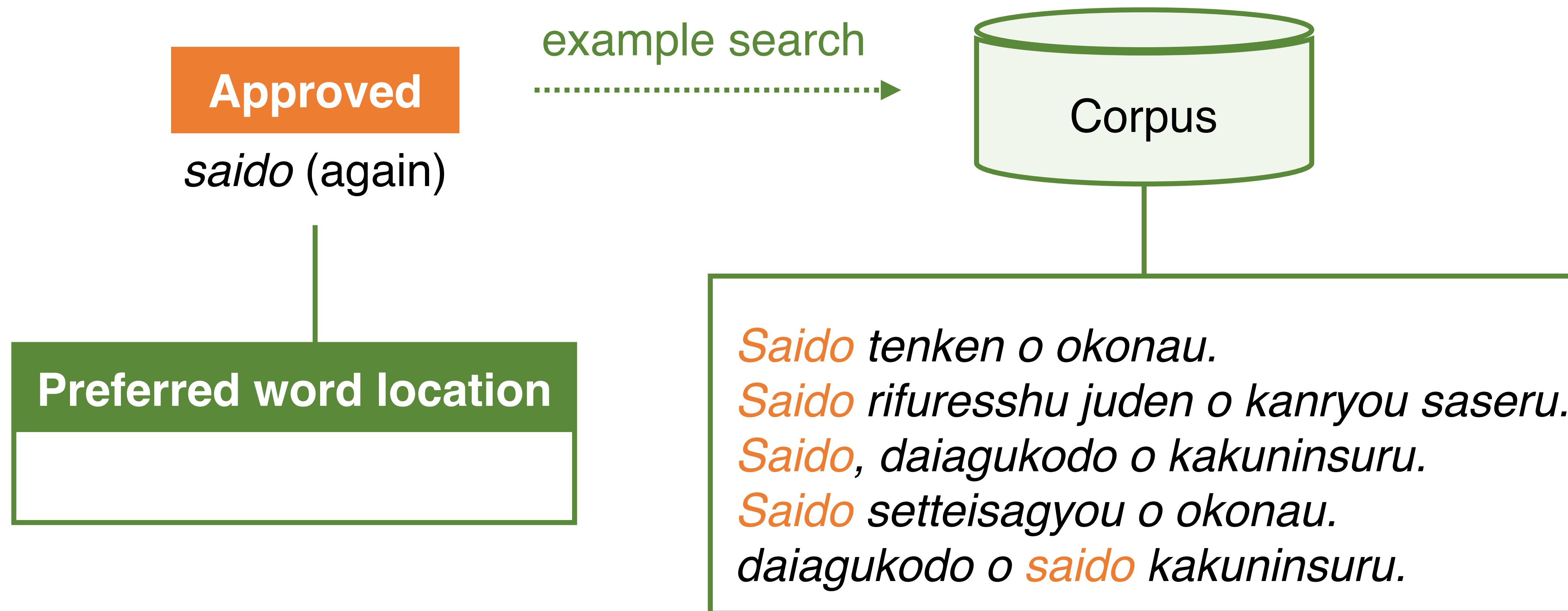
*saido* (again)

**Preferred word location**



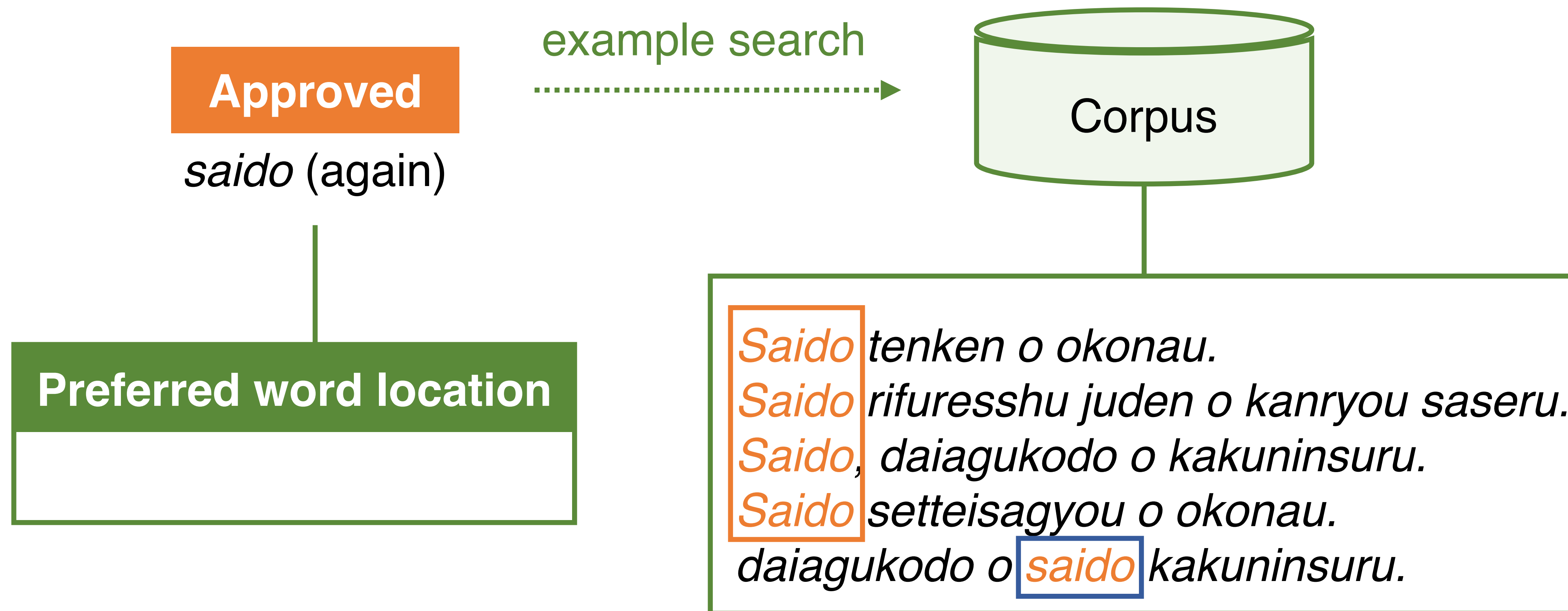
# Building Procedure

## Definition of preferred word location



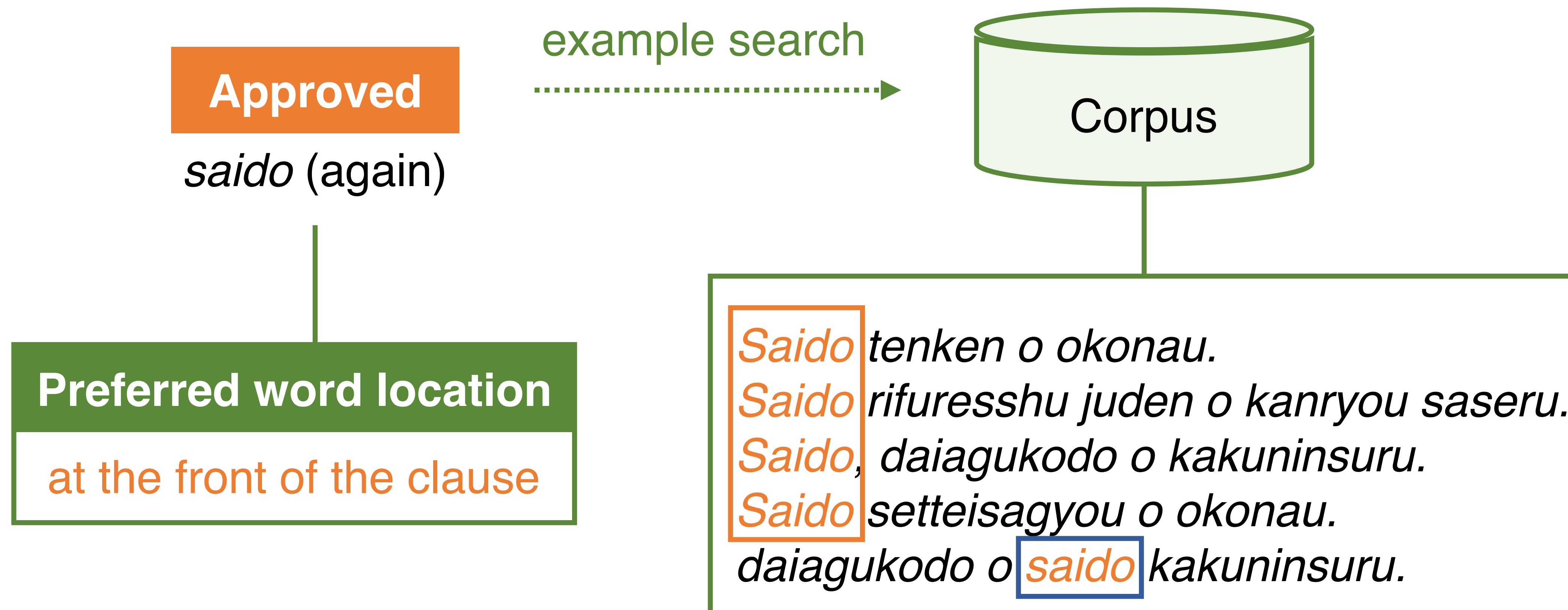
# Building Procedure

## Definition of preferred word location



# Building Procedure

## Definition of preferred word location



# Results

## ■ Fixed types

Category		Example entry	Approved		Unapproved	
Level 1	Level 2		# Type	# Token	# Type	# Token
manner		<i>yukkurito</i> → <i>yukkuri</i> (slowly)	110	21,779	12	660
degree		<i>ooyoso</i> → <i>hobo</i> (almost)	17	2,431	3	92
aspect	proximity	<i>tadachini</i> → <i>suguni</i> (immediately)	2	381	3	241
	continuation	<i>sukoshizutsu</i> → <i>jojoni</i> (gradually)	9	2,937	2	119
	repetition	<i>futatabi</i> → <i>saido</i> (again)	2	3,971	1	143
	order	<i>sonoatoni</i> → <i>sonogo</i> (then)	21	5,145	5	104
	frequency	<i>jouji</i> → <i>tsuneni</i> (always)	6	2,100	2	37
emphasis		<i>kanarazu</i> (surely)	3	18,475	1	132
others		<i>sorezore</i> (respectively)	65	6,183	9	185
<b>Total</b>			235	63,402	38	1,713

97.4%

2.6%



# Results

## ■ Fixed types

Category		Example entry	Approved		Unapproved	
Level 1	Level 2		# Type	# Token	# Type	# Token
manner		<i>yukkurito</i> → <i>yukkuri</i> (slowly)	110	21,779	12	660
degree		<i>ooyoso</i> → <i>hobo</i> (almost)	17	2,431	3	92
aspect	proximity	<i>tadachini</i> → <i>suguni</i> (immediately)	2	381	3	241
	continuation	<i>sukoshizutsu</i> → <i>jojoni</i> (gradually)	9	2,937	2	119
	repetition	<i>futatabi</i> → <i>saido</i> (again)	2	3,971	1	143
	order	<i>sonoatoni</i> → <i>sonogo</i> (then)	21	5,145	5	104
	frequency	<i>jouji</i> → <i>tsuneni</i> (always)	6	2,100	2	37
emphasis		<i>kanarazu</i> (surely)	3	18,475	1	132
others		<i>sorezore</i> (respectively)	65	6,183	9	185
<b>Total</b>			235	63,402	38	1,713

97.4%

2.6%

# Results

## ■ Fixed types

Category		Example entry	Approved		Unapproved	
Level 1	Level 2		# Type	# Token	# Type	# Token
manner		<i>yukkurito</i> → <i>yukkuri</i> (slowly)	110	21,779	12	660
degree		<i>ooyoso</i> → <i>hobo</i> (almost)	17	2,431	3	92
aspect	proximity	<i>tadachini</i> → <i>suguni</i> (immediately)	2	381	3	241
	continuation	<i>sukoshizutsu</i> → <i>jojoni</i> (gradually)	9	2,937	2	119
	repetition	<i>futatabi</i> → <i>saido</i> (again)	2	3,971	1	143
	order	<i>sonoatoni</i> → <i>sonogo</i> (then)	21	5,145	5	104
	frequency	<i>jouji</i> → <i>tsuneni</i> (always)	6	2,100	2	37
emphasis		<i>kanarazu</i> (surely)	3	18,475	1	132
others		<i>sorezore</i> (respectively)	65	6,183	9	185
<b>Total</b>			<b>235</b>	<b>63,402</b>	<b>38</b>	<b>1,713</b>

97.4%

2.6%

# Results

## ■ Fixed types

Category		Example entry	Approved		Unapproved	
Level 1	Level 2		# Type	# Token	# Type	# Token
manner		<i>yukkurito</i> → <i>yukkuri</i> (slowly)	110	21,779	12	660
degree		<i>ooyoso</i> → <i>hobo</i> (almost)	17	2,431	3	92
aspect	proximity	<i>tadachini</i> → <i>suguni</i> (immediately)	2	381	3	241
	continuation	<i>sukoshizutsu</i> → <i>jojoni</i> (gradually)	9	2,937	2	119
	repetition	<i>futatabi</i> → <i>saido</i> (again)	2	3,971	1	143
	order	<i>sonoatoni</i> → <i>sonogo</i> (then)	21	5,145	5	104
	frequency	<i>jouji</i> → <i>tsuneni</i> (always)	6	2,100	2	37
emphasis		<i>kanarazu</i> (surely)	3	18,475	1	132
others		<i>sorezore</i> (respectively)	65	6,183	9	185
<b>Total</b>			<b>235</b>	<b>63,402</b>	<b>38</b>	<b>1,713</b>
				97.4%		2.6%

# Results

## ■ Fixed types

Category		Example entry	Approved		Unapproved	
Level 1	Level 2		# Type	# Token	# Type	# Token
manner		<i>yukkurito</i> → <i>yukkuri</i> (slowly)	110	21,779	12	660
degree		<i>ooyoso</i> → <i>hobo</i> (almost)	17	2,431	3	92
aspect	proximity	<i>tadachini</i> → <i>suguni</i> (immediately)	2	381	3	241
	continuation	<i>sukoshizutsu</i> → <i>jojoni</i> (gradually)	9	2,937	2	119
	repetition	<i>futatabi</i> → <i>saido</i> (again)	2	3,971	1	143
	order	<i>sonoatoni</i> → <i>sonogo</i> (then)	21	5,145	5	104
	frequency	<i>jouji</i> → <i>tsuneni</i> (always)	6	2,100	2	37
emphasis		<i>kanarazu</i> (surely)	3	18,475	1	132
others		<i>sorezore</i> (respectively)	65	6,183	9	185
<b>Total</b>			235	63,402	38	1,713

97.4%

2.6%

# Results

## Variable Types

Category Level 1	Example entry	Approved		Unapproved	
		# Type	# Token	# Type	# Token
manner	[X] to douyou ni (in the same way as [X]), [X] to issho ni (with [X])	3	605	0	0
quantity	[X] byou ijou → sukunakutomo [X] byoukan (for at least [X] seconds), [X] fun → [X] funkan (for [X] minutes)	37	16,953	54	6,601
time	[X] go ni wa → [X] go wa (after [X]), [X] ji nado de → [X] ji nado ni (when [X])	31	54,077	10	118
order	[X] kara jun ni (from [X] in order), [X] no [Y] jun ni (in [Y] order of the [X])	4	2,341	0	0
<b>Total</b>		75	73,976	64	6,719

91.7%

8.3%

# Results

## ■ Variable Types

Category Level 1	Example entry	Approved		Unapproved	
		# Type	# Token	# Type	# Token
manner	[X] to douyou ni (in the same way as [X]), [X] to issho ni (with [X])	3	605	0	0
quantity	[X] byou ijou → sukunakutomo [X] byoukan (for at least [X] seconds), [X] fun → [X] funkan (for [X] minutes)	37	16,953	54	6,601
time	[X] go ni wa → [X] go wa (after [X]), [X] ji nado de → [X] ji nado ni (when [X])	31	54,077	10	118
order	[X] kara jun ni (from [X] in order), [X] no [Y] jun ni (in [Y] order of the [X])	4	2,341	0	0
<b>Total</b>		<b>75</b>	<b>73,976</b>	<b>64</b>	<b>6,719</b>

91.7%

8.3%

# Results

## Variable Types

Category Level 1	Example entry	Approved		Unapproved	
		# Type	# Token	# Type	# Token
manner	[X] to douyou ni (in the same way as [X]), [X] to issho ni (with [X])	3	605	0	0
quantity	[X] byou ijou → sukunakutomo [X] byoukan (for at least [X] seconds), [X] fun → [X] funkan (for [X] minutes)	37	16,953	54	6,601
time	[X] go ni wa → [X] go wa (after [X]), [X] ji nado de → [X] ji nado ni (when [X])	31	54,077	10	118
order	[X] kara jun ni (from [X] in order), [X] no [Y] jun ni (in [Y] order of the [X])	4	2,341	0	0
<b>Total</b>		75	73,976	64	6,719

91.7%

8.3%

# Lexicon Building: Sentence-ending Functions

the clip A    from the vehicle    be removed    **have**    when    surely    with a new one    replace    **be sure to**  
[クリップA を] [車両 から] 取りはずし **た** 場合は、必ず [新品 に] 交換する **こと**  
*kurippu A o    sharyou kara    torihazushi    ta    baai wa    kanarazu    shinpin ni    koukansuru    koto*

- case frame
- case element
- adverbial modifier
- **sentence-ending function**
- connective

sequences of functional words  
attached to sentence-ending main verbs



# Sentence-ending Functions

- The sentence-ending functions are word sequences that can be attached to the main verbs for adding various functional information (Kato et al. 2020)

display

表示さ れ ない 場合 が ある (may not be displayed)

*hyoujisa- re nai baai ga aru*

# Sentence-ending Functions

- The sentence-ending functions are word sequences that can be attached to the main verbs for adding various functional information (Kato et al. 2020)

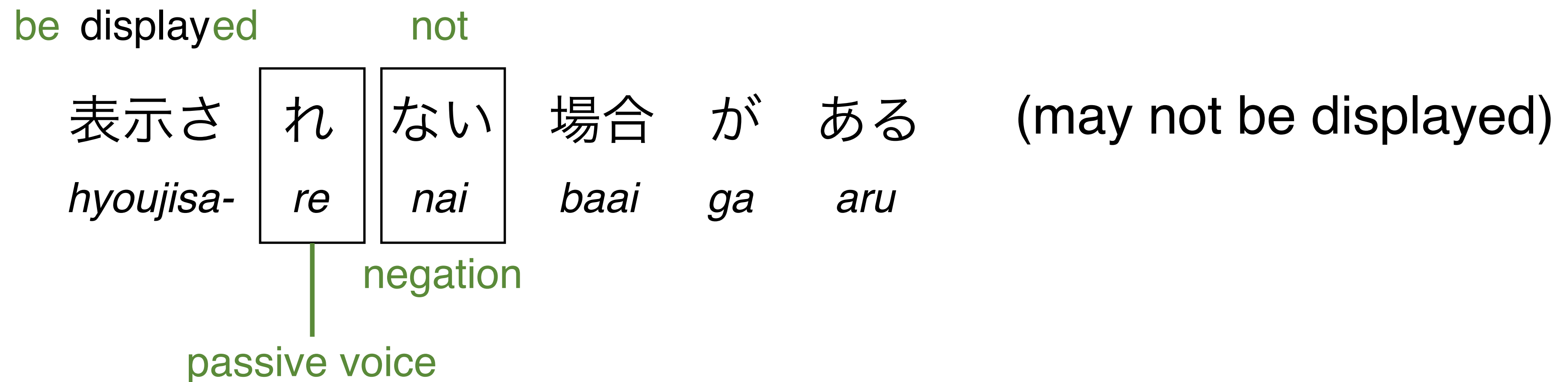
be displayed

表示さ れ ない 場合 が ある (may not be displayed)  
*hyoujisa- re nai baai ga aru*

passive voice

# Sentence-ending Functions

- The sentence-ending functions are word sequences that can be attached to the main verbs for adding various functional information (Kato et al. 2020)



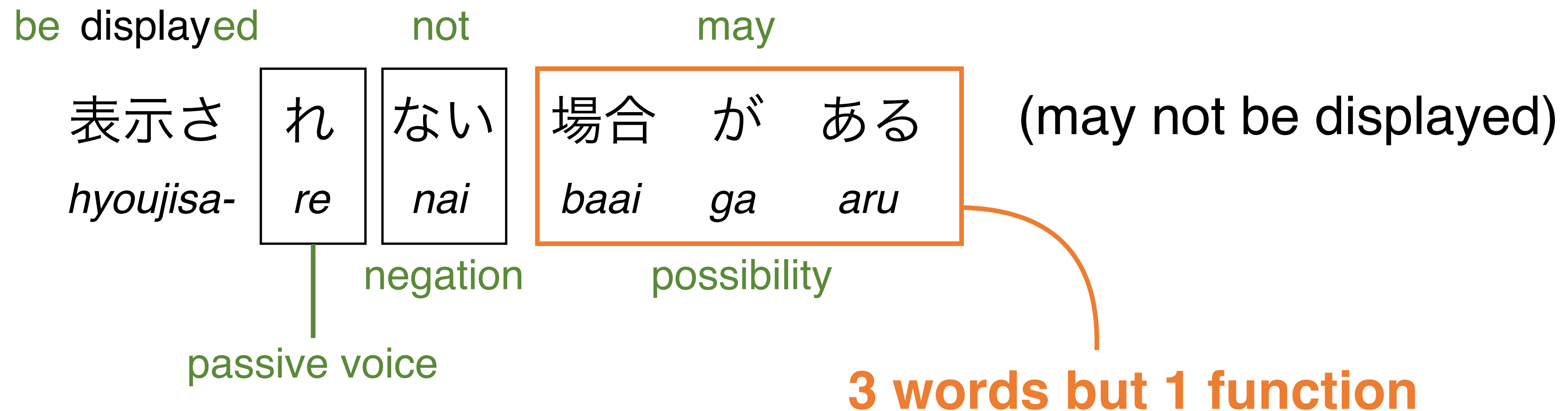
# Sentence-ending Functions

- The sentence-ending functions are word sequences that can be attached to the main verbs for adding various functional information (Kato et al. 2020)



# Sentence-ending Functions

- The sentence-ending functions are word sequences that can be attached to the main verbs for adding various functional information (Kato et al. 2020)



# Building Procedure

(Trans.: The measurement value may not be displayed correctly.)

*tadashii sokuteichi ga hyoujisarenai baai ga aru*

(Trans.: Check the fuses.)

*hyuzu o tenkensuru koto*

(Trans.: The history may be cleared.)

*Rireki ga kieteshimau koto ga aru*

# Building Procedure

(Trans.: The measurement value may not be displayed correctly.)

*tadashii sokuteichi ga hyoujisarenai baai ga aru*

(Trans.: Check the fuses.)

*hyuzu o tenkensuru koto*

(Trans.: The history may be cleared.)

*Rireki ga kieteshimau koto ga aru*

**extract 150 most frequent  
sentence-ending patterns**

# Building Procedure

(Trans.: The measurement value may not be displayed correctly.)

*tadashii sokuteichi ga hyoujisarenai baai ga aru*

(Trans.: Check the fuses.)

*hyuzu o tenkensuru koto*

(Trans.: The history may be cleared.)

*Rireki ga kieteshimau koto ga aru*

**extract 150 most frequent sentence-ending patterns**

**decompose them into single functions**

*re* *nai* *baai ga aru*

*koto*

*te-shimau* *koto ga aru*

⋮



# Building Procedure

(Trans.: The measurement value may not be displayed correctly.)

*tadashii sokuteichi ga hyoujisarenai baai ga aru*

(Trans.: Check the fuses.)

*hyuzu o tenkensuru koto*

(Trans.: The history may be cleared.)

*Rireki ga kieteshimau koto ga aru*

extract 150 most frequent sentence-ending patterns

decompose them into single functions

grouping

*re* *nai*

*koto*

*te-shimau*

⋮

*baai ga aru*

*koto ga aru*

# Building Procedure

(Trans.: The measurement value may not be displayed correctly.)

*tadashii sokuteichi ga hyoujisarenai baai ga aru*

(Trans.: Check the fuses.)

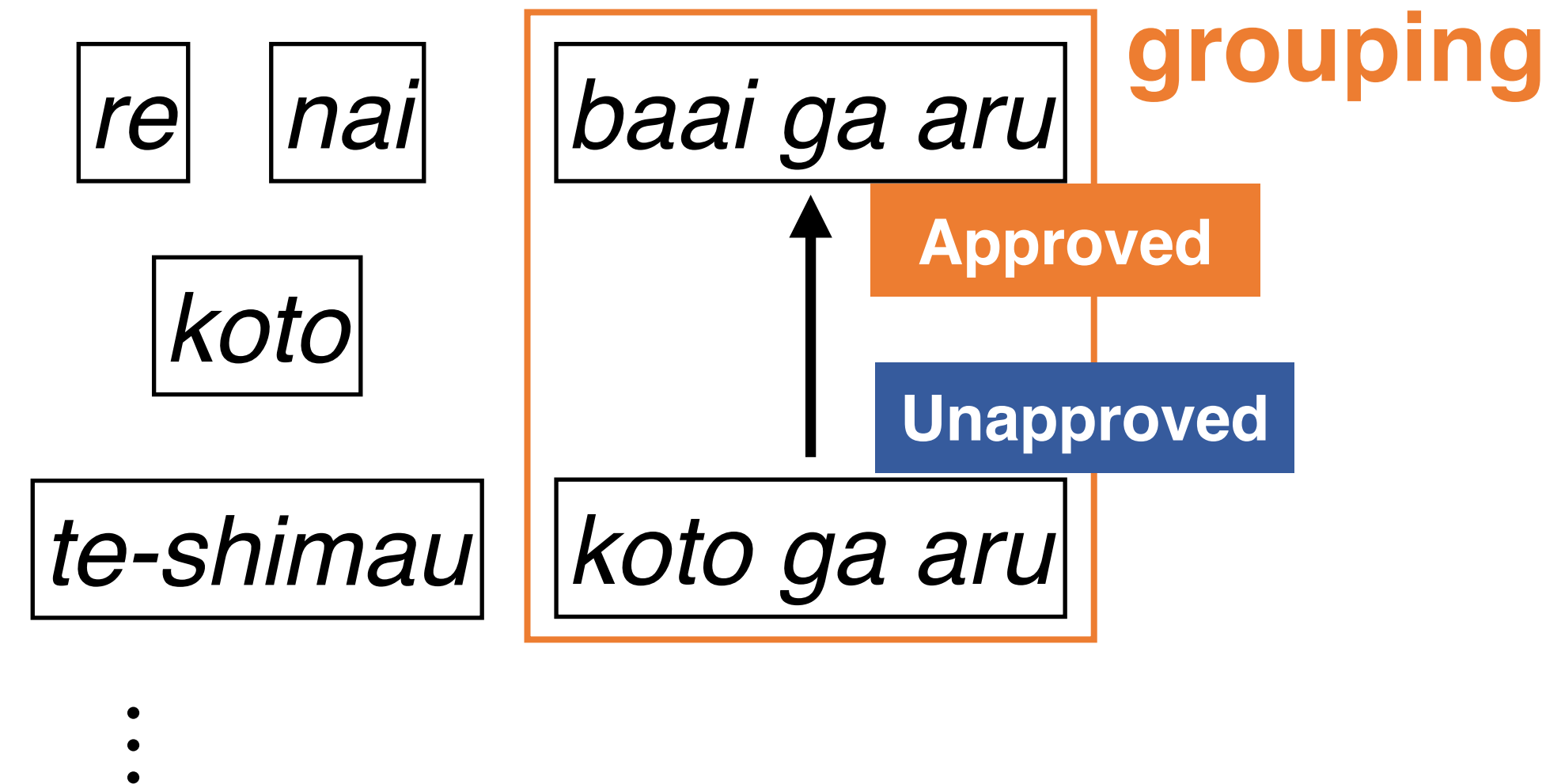
*hyuzu o tenkensuru koto*

(Trans.: The history may be cleared.)

*Rireki ga kieteshimau koto ga aru*

extract 150 most frequent sentence-ending patterns

decompose them into single functions







# Example: Possibility Modality

(Trans.: waiting time **may** be required)

*Machijikan ga hasseisuru baai ga aru*  
場合がある

(Trans.: a DTC **may** be output)

*Daiagu o shutsuryokusuru kanousei ga aru*  
可能性がある

(Trans.: this **may** cause a malfunction)

*Fuguai ga hasseisuru osore ga aru*  
おそれがある

# Example: Possibility Modality

(Trans.: waiting time **may** be required)

*Machijikan ga hasseisuru baai ga aru*  
場合がある

there is a case

(Trans.: a DTC **may** be output)

*Daiagu o shutsuryokusuru kanousei ga aru*  
可能性がある

there is a possibility

(Trans.: this **may** cause a malfunction)

*Fuguai ga hasseisuru osore ga aru*  
おそれがある

there is a risk

# Lexicon Building: Connectives

the clip A    from the vehicle    be removed    have    **when**    surely    with a new one    replace    be sure to  
[クリップA を] [車両 から] 取りはずし た **場合は**、必ず [新品 に] 交換する こと  
*kurippu A o    sharyou kara    torihazushi    ta    baai wa    kanarazu    shinpin ni    koukansuru    koto*

- case frame
- case element
- adverbial modifier
- sentence-ending function
- **connective**

conjunctions and conjunctive phrases  
that indicate inter-clause or inter-sentence relationships

# Types of Connectives

## ■ Clause-beginning type (inter-sentence connection)

(Trans.: **Therefore**, make sure to read the notices before proceeding with work.)

*Konotame, sagyoumae ni chuijiko o kakuninsuru*

このため

## ■ Clause-ending type (inter-clause connection)

(Trans.: **When** the clip A has been removed from the vehicle, be sure to replace with a new one)

*kurippu A o sharyou kara torihazushi ta baai wa,  
kanarazu shinpin ni koukansuru koto*      場合は



# Building Procedure

- Clause-beginning type
  - Conjunctions and conjunctive adverbs
  - Heuristic rules with Juman++ (morphological analyser)
- Clause-ending type
  - Heuristic rules with Juman++ and KNP (syntactic analyser)

## Rules for coordinate and subordinate clauses

(Trans.: **When** the battery voltage is applied to each terminal, check the illumination condition of the light )

*Kakutanshi ni batteridenatsu o kuwaeta toki, ranpu no tentoujoutai o kakuninsuru*

## Rules for attributive clauses

(Trans.: **Before** performing the troubleshooting, check the fuses for circuits related to this system )

*Toraburushuto o jisshisuru mae ni, kono kairo no hyuzu o tenkensuru koto.*

# Building Procedure

- Clause-beginning type
  - Conjunctions and conjunctive adverbs
  - Heuristic rules with Juman++ (morphological analyser)
- Clause-ending type
  - Heuristic rules with Juman++ and KNP (syntactic analyser)

## Rules for coordinate and subordinate clauses

(Trans.: **When** the battery voltage is applied to each terminal, check the illumination condition of the light )

*Kakutanshi ni batteridenatsu o kuwaeta toki, ranpu no tentoujoutai o kakuninsuru*

subordinate clause

## Rules for attributive clauses

(Trans.: **Before** performing the troubleshooting, check the fuses for circuits related to this system )

*Toraburushuto o jisshisuru mae ni, kono kairo no hyuzu o tenkensuru koto.*

# Building Procedure

- Clause-beginning type
  - Conjunctions and conjunctive adverbs
  - Heuristic rules with Juman++ (morphological analyser)
- Clause-ending type
  - Heuristic rules with Juman++ and KNP (syntactic analyser)

## Rules for coordinate and subordinate clauses

(Trans.: **When** the battery voltage is applied to each terminal, check the illumination condition of the light )

*Kakutanshi ni batteridenatsu o kuwaeta toki, ranpu no tentoujoutai o kakuninsuru*

subordinate clause

*toki (when)*

## Rules for attributive clauses

(Trans.: **Before** performing the troubleshooting, check the fuses for circuits related to this system )

*Toraburushuto o jisshisuru mae ni, kono kairo no hyuzu o tenkensuru koto.*

# Building Procedure

- Clause-beginning type
  - Conjunctions and conjunctive adverbs
  - Heuristic rules with Juman++ (morphological analyser)
- Clause-ending type
  - Heuristic rules with Juman++ and KNP (syntactic analyser)

## Rules for coordinate and subordinate clauses

(Trans.: **When** the battery voltage is applied to each terminal, check the illumination condition of the light )

*Kakutanshi ni batteridenatsu o kuwaeta toki, ranpu no tentoujoutai o kakuninsuru*

subordinate clause

*toki (when)*

## Rules for attributive clauses

(Trans.: **Before** performing the troubleshooting, check the fuses for circuits related to this system )

*Toraburushuto o jisshisuru mae ni, kono kairo no hyuzu o tenkensuru koto.*

attributive clause

# Building Procedure

- Clause-beginning type
  - Conjunctions and conjunctive adverbs
  - Heuristic rules with Juman++ (morphological analyser)
- Clause-ending type
  - Heuristic rules with Juman++ and KNP (syntactic analyser)

## Rules for coordinate and subordinate clauses

(Trans.: **When** the battery voltage is applied to each terminal, check the illumination condition of the light )

*Kakutanshi ni batteridenatsu o kuwaeta toki, ranpu no tentoujoutai o kakuninsuru*

subordinate clause

*toki (when)*

## Rules for attributive clauses

(Trans.: **Before** performing the troubleshooting, check the fuses for circuits related to this system )

*Toraburushuto o jisshisuru mae ni, kono kairo no hyuzu o tenkensuru koto.*

attributive clause

parent

# Building Procedure

- Clause-beginning type
  - Conjunctions and conjunctive adverbs
  - Heuristic rules with Juman++ (morphological analyser)
- Clause-ending type
  - Heuristic rules with Juman++ and KNP (syntactic analyser)

## Rules for coordinate and subordinate clauses

(Trans.: **When** the battery voltage is applied to each terminal, check the illumination condition of the light )

*Kakutanshi ni batteridenatsu o kuwaeta toki, ranpu no tentoujoutai o kakuninsuru*

subordinate clause

*toki (when)*

## Rules for attributive clauses

(Trans.: **Before** performing the troubleshooting, check the fuses for circuits related to this system )

*Toraburushuto o jisshisuru mae ni, kono kairo no hyuzu o tenkensuru koto.*

attributive clause

parent

*mae ni (before)*

# Results: Clause-beginning Connectives

Category	Example	Approved		Unapproved	
		# Type	# Token	# Type	# Token
resultative	<i>shitagatte</i> (kana) → <i>shitagatte</i> (kanji) (therefore), <i>yotte</i> → <i>shitagatte</i> (kanji) (therefore), <i>konotame</i> (for this reason)	4	625	2	135
adversative	<i>gyakuni</i> (conversely), <i>shikashi</i> (but)	2	96	0	0
parallel	<i>mata</i> (kanji) → <i>mata</i> (kana) (also), <i>katsu</i> (kana) → <i>katsu</i> (kanji) (besides), <i>sarani</i> (in addition)	6	4,915	2	23
contrast	<i>mataha</i> (kanji) → <i>mataha</i> (kana) (or), <i>aruiha</i> → <i>mataha</i> (kana) (or), <i>moshikuha</i> → <i>mataha</i> (kana) (or)	1	124	3	40
paraphrase	<i>tsumari</i> (in other words)	1	49	0	0
example	<i>tatoeba</i> (kana) → <i>tatoeba</i> (kanji) (for example)	1	31	1	2
addition	<i>tadashi</i> (kanji) → <i>tadashi</i> (kana) (however), <i>nao</i> (in addition)	2	609	1	398
	<b>Total</b>	17	6,449	9	598

91.5%

8.5%

# Results: Clause-beginning Connectives

Category	Example	Approved		Unapproved	
		# Type	# Token	# Type	# Token
resultative	<i>shitagatte</i> (kana) → <i>shitagatte</i> (kanji) (therefore), <i>yotte</i> → <i>shitagatte</i> (kanji) (therefore), <i>konotame</i> (for this reason)	4	625	2	135
adversative	<i>gyakuni</i> (conversely), <i>shikashi</i> (but)	2	96	0	0
parallel	<i>mata</i> (kanji) → <i>mata</i> (kana) (also), <i>katsu</i> (kana) → <i>katsu</i> (kanji) (besides), <i>sarani</i> (in addition)	6	4,915	2	23
contrast	<i>mataha</i> (kanji) → <i>mataha</i> (kana) (or), <i>aruiha</i> → <i>mataha</i> (kana) (or), <i>moshikuha</i> → <i>mataha</i> (kana) (or)	1	124	3	40
paraphrase	<i>tsumari</i> (in other words)	1	49	0	0
example	<i>tatoeba</i> (kana) → <i>tatoeba</i> (kanji) (for example)	1	31	1	2
addition	<i>tadashi</i> (kanji) → <i>tadashi</i> (kana) (however), <i>nao</i> (in addition)	2	609	1	398
	<b>Total</b>	17	6,449	9	598
			91.5%		8.5%



# Results: Clause-beginning Connectives

Category	Example	Approved		Unapproved	
		# Type	# Token	# Type	# Token
resultative	<i>shitagatte</i> (kana) → <i>shitagatte</i> (kanji) (therefore), <i>yotte</i> → <i>shitagatte</i> (kanji) (therefore), <i>konotame</i> (for this reason)	4	625	2	135
adversative	<i>gyakuni</i> (conversely), <i>shikashi</i> (but)	2	96	0	0
parallel	<i>mata</i> (kanji) → <i>mata</i> (kana) (also), <i>katsu</i> (kana) → <i>katsu</i> (kanji) (besides), <i>sarani</i> (in addition)	6	4,915	2	23
contrast	<i>mataha</i> (kanji) → <i>mataha</i> (kana) (or), <i>aruiha</i> → <i>mataha</i> (kana) (or), <i>moshikuha</i> → <i>mataha</i> (kana) (or)	1	124	3	40
paraphrase	<i>tsumari</i> (in other words)	1	49	0	0
example	<i>tatoeba</i> (kana) → <i>tatoeba</i> (kanji) (for example)	1	31	1	2
addition	<i>tadashi</i> (kanji) → <i>tadashi</i> (kana) (however), <i>nao</i> (in addition)	2	609	1	398
	<b>Total</b>	17	6,449	9	598

91.5%

8.5%

# Results: Clause-ending Connectives

Category		Example entry	Approved		Unapproved	
Level 1	Level 2		# Type	# Token	# Type	# Token
coordinate	resultative	<i>V-te</i> → <i>V</i> (continuative form)	1	218,165	1	199,945
	illustration	<i>tari, nado</i>	2	6,737	0	0
	contradictory	<i>ga</i>	1	4,912	0	0
	cumulation	<i>ue</i>	1	17	0	0
adverbial	time	<i>toki</i> (kanji)/ <i>tokiwa</i> (kana/kanji) → <i>toki</i> (kana)	13	68,661	22	23,814
	condition	<i>baaini</i> → <i>baai</i>	5	113,518	17	8,258
	method	<i>V</i> (continuative form) → <i>V-te/kotode</i>	2	7,019	1	0
	attendant circumstances	<i>tsutsu</i> → <i>nagara/mama</i>	3	12,934	1	27
	state	<i>V</i> (continuative form) → <i>V-te/joutaide</i>	2	7364	1	0
	purpose	<i>tame</i> (kanji)/ <i>tameni</i> (kana/kanji) → <i>tame</i> (kana)	3	10,083	7	3,375
	reason	<i>tame</i> (kana/kanji) → <i>node</i>	1	3,100	3	37,703
	others	<i>noni, teido</i> → <i>teidoni</i>	4	479	3	305
		<b>Total</b>	37	452,989	43	273,427
				62.4%		37.6%

# Results: Clause-ending Connectives

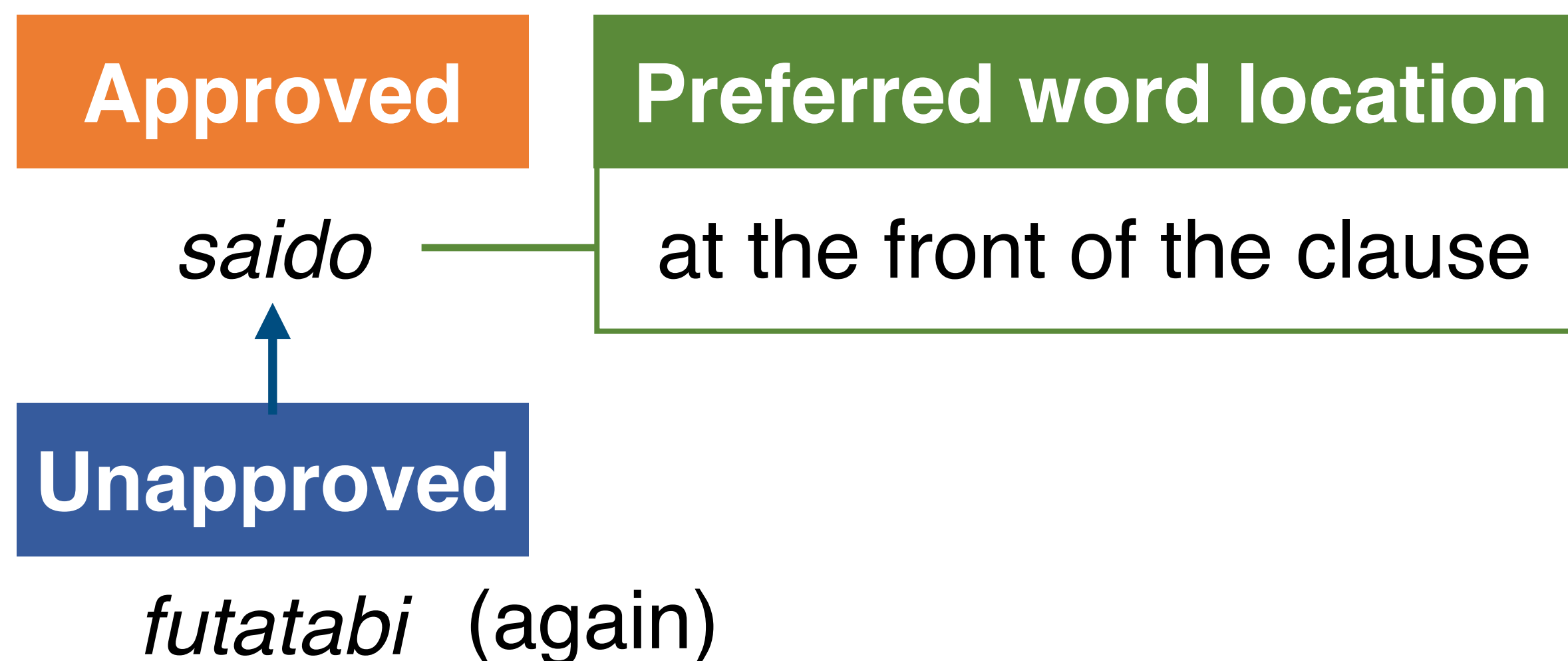
Category		Example entry	Approved		Unapproved	
Level 1	Level 2		# Type	# Token	# Type	# Token
coordinate	resultative	<i>V-te</i> → <i>V</i> (continuative form)	1	218,165	1	199,945
	illustration	<i>tari, nado</i>	2	6,737	0	0
	contradictory	<i>ga</i>	1	4,912	0	0
	cumulation	<i>ue</i>	1	17	0	0
adverbial	time	<i>toki</i> (kanji)/ <i>tokiwa</i> (kana/kanji) → <i>toki</i> (kana)	13	68,661	22	23,814
	condition	<i>baaini</i> → <i>baai</i>	5	113,518	17	8,258
	method	<i>V</i> (continuative form) → <i>V-te/kotode</i>	2	7,019	1	0
	attendant circumstances	<i>tsutsu</i> → <i>nagara/mama</i>	3	12,934	1	27
	state	<i>V</i> (continuative form) → <i>V-te/joutaide</i>	2	7364	1	0
	purpose	<i>tame</i> (kanji)/ <i>tameni</i> (kana/kanji) → <i>tame</i> (kana)	3	10,083	7	3,375
	reason	<i>tame</i> (kana/kanji) → <i>node</i>	1	3,100	3	37,703
	others	<i>noni, teido</i> → <i>teidoni</i>	4	479	3	305
		<b>Total</b>	<b>37</b>	<b>452,989</b>	<b>43</b>	<b>273,427</b>
				62.4%		37.6%

# Results: Clause-ending Connectives

Category		Example entry	Approved		Unapproved	
Level 1	Level 2		# Type	# Token	# Type	# Token
coordinate	resultative	<i>V-te</i> → <i>V</i> (continuative form)	1	218,165	1	199,945
	illustration	<i>tari, nado</i>	2	6,737	0	0
	contradictory	<i>ga</i>	1	4,912	0	0
	cumulation	<i>ue</i>	1	17	0	0
adverbial	time	<i>toki</i> (kanji)/ <i>tokiwa</i> (kana/kanji) → <i>toki</i> (kana)	13	68,661	22	23,814
	condition	<i>baaini</i> → <i>baai</i>	5	113,518	17	8,258
	methond	<i>V</i> (continuative form) → <i>V-te/kotode</i>	2	7,019	1	0
	attendant circumstances	<i>tsutsu</i> → <i>nagara/mama</i>	3	12,934	1	27
	state	<i>V</i> (continuative form) → <i>V-te/joutaide</i>	2	7364	1	0
	purpose	<i>tame</i> (kanji)/ <i>tameni</i> (kana/kanji) → <i>tame</i> (kana)	3	10,083	7	3,375
	reason	<i>tame</i> (kana/kanji) → <i>node</i>	1	3,100	3	37,703
	others	<i>noni, teido</i> → <i>teidoni</i>	4	479	3	305
		<b>Total</b>	<b>37</b>	<b>452,989</b>	<b>43</b>	<b>273,427</b>
				62.4%		37.6%

# Conclusion and Outlook

- We built a Japanese controlled language for the automotive domain
  - based on the analysis of a large-scale text corpus
  - consisting of 5 lexical components
    - ◆ case frame
    - ◆ case element
    - ◆ **adverbial modifier**
    - ◆ **sentence-ending function**
    - ◆ **connective**
- We will fully develop the writing assistant tool and evaluate its usability



Thank you!