



GENERATING **CONTRASTIVE** REFERRING EXPRESSIONS

Martín Villalba, Christoph Teichmann,
and Alexander Koller







Intended target

The yellow cup to the right of the jar





Intended target



The yellow cup to the right of the jar



User resolved target





No, not that one

No, the yellow cup to the right
of the jar

No, the mug with squares



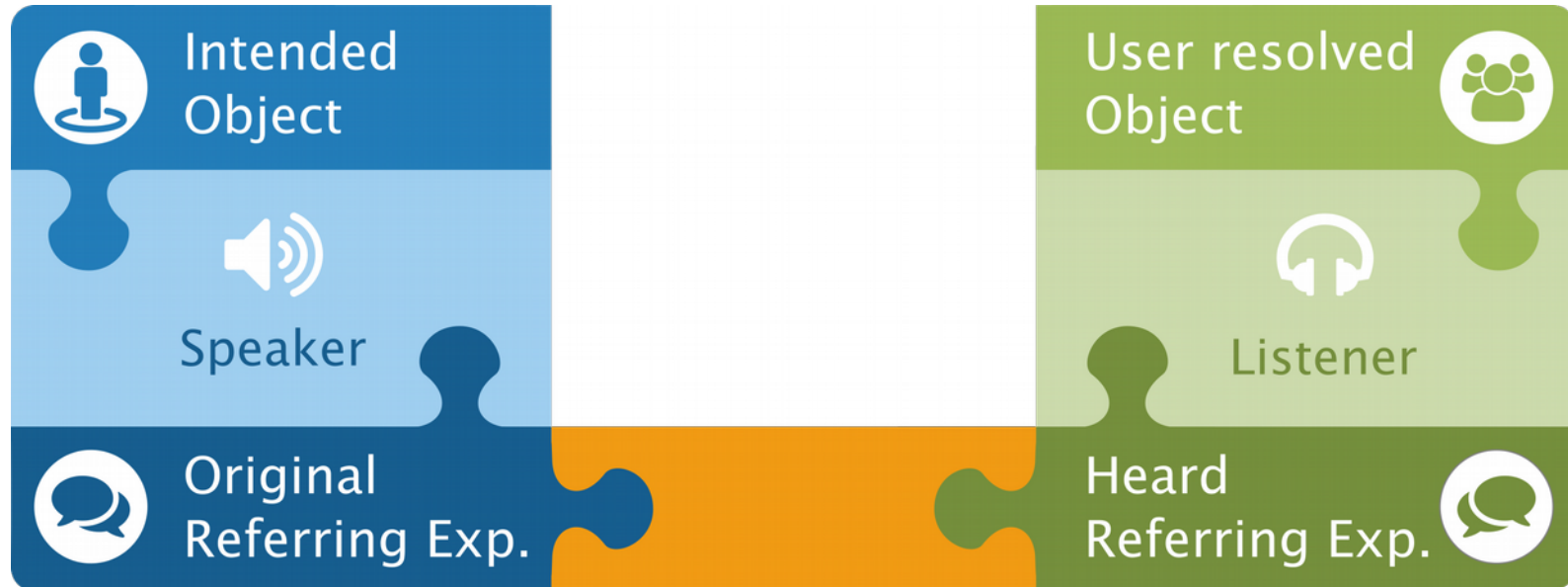
No, not that one

No, the yellow cup to the right
of the jar

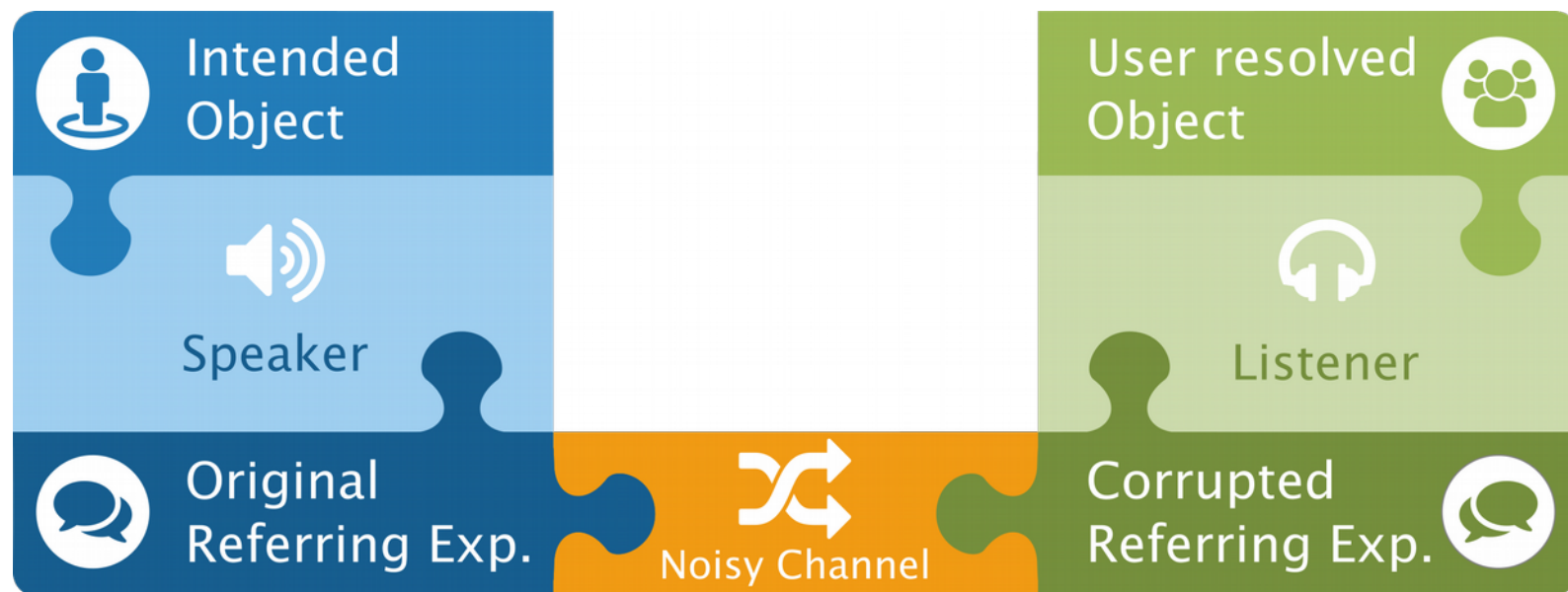
No, the mug with squares

No, the **YELLOW** cup
to the right of the jar

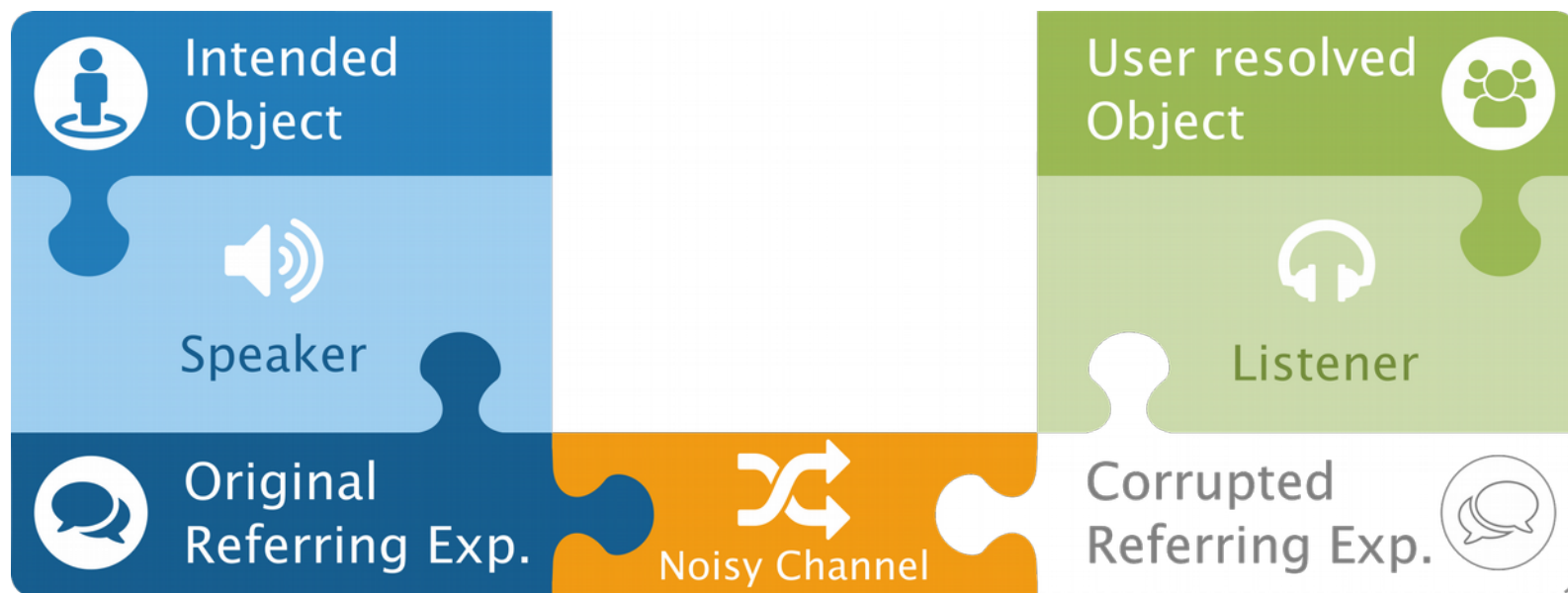
Communication model



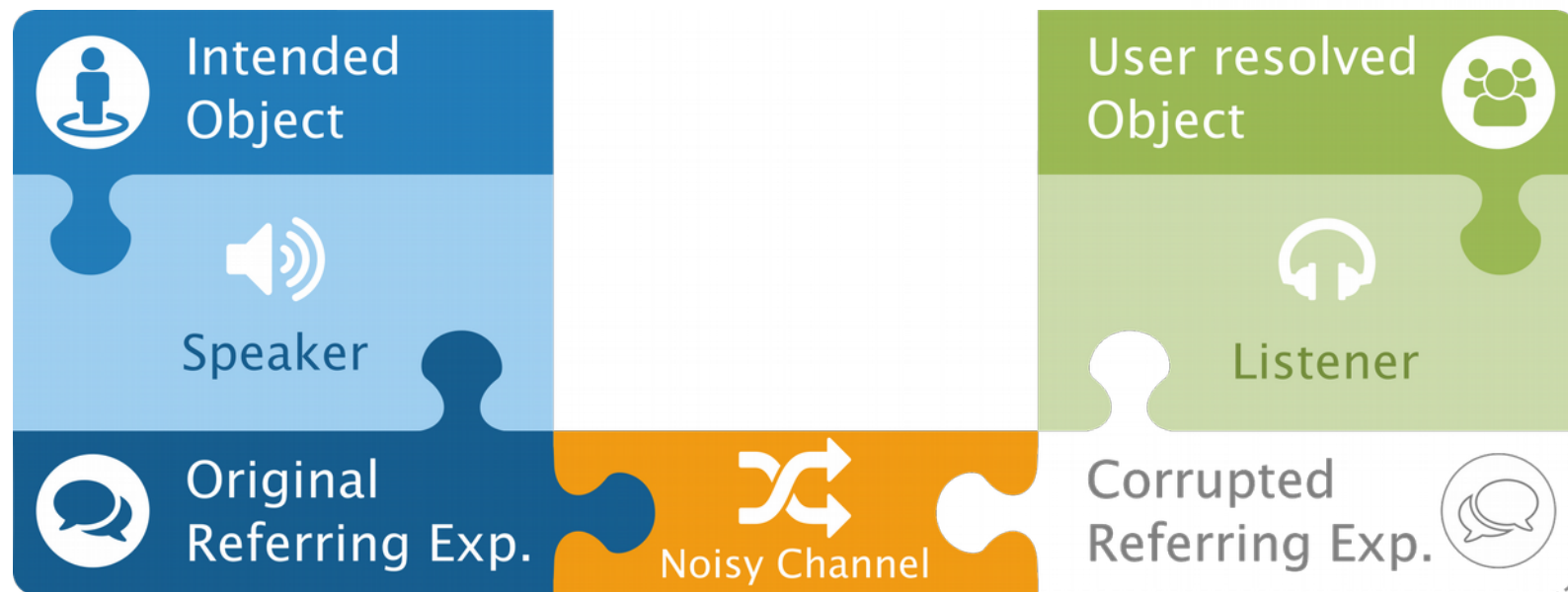
Corruption model



Corruption model



Corruption model



Problem 1: Find the corrupted RE and the corruption from the Noisy Channel

Problem 2: Generate a contrastive RE based on this information

01

PART I

Finding the corrupted RE

How we do it

Find the RE that minimizes the
Levenshtein edit distance over words
between the *Original* RE and the *Heard* RE,
while referring to the User Resolved Object

Step 1: Find all possible REs

Our approach requires a Context Free Grammar G such that $L(G)$ is the set of all REs referring to the User Resolved Object

We use the algorithm of *Engonopoulos and Koller* to find it efficiently.

Step 1: Find all possible REs



The cup to the right of the yellow cup

The cup to the right of the cup

The Berlin cup

The white cup

The cup to the left of the water boiler

The cup above the folder with
a red label

The cup in front of the green book

The cup in front of the cup

The white cup to the right of the jar

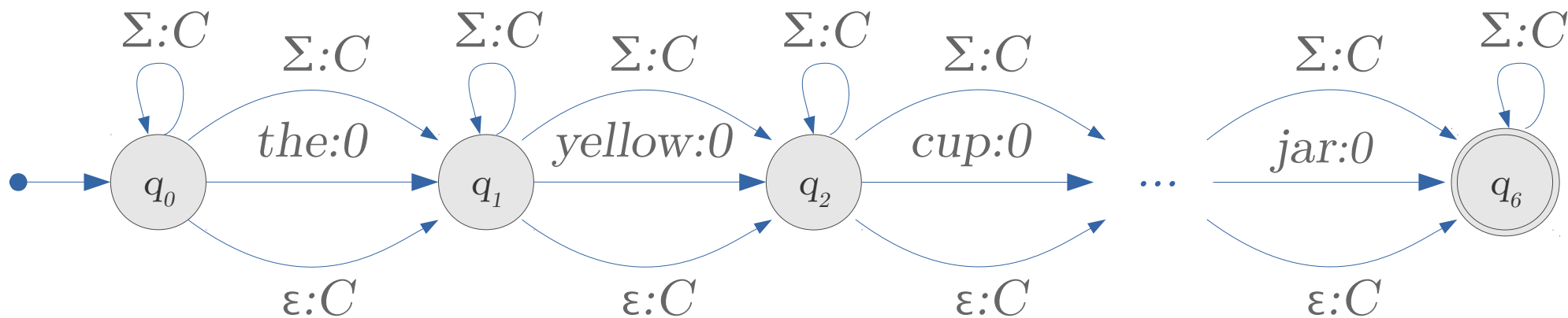
The cup without squares

The cup to the right of the cup to the
right of the jar

...

Step 2: Corruption distances

We derive a weighted finite-state automaton F , where each run of the automaton on a string w is an edit sequence transforming our Original RE into w





Step 3: Compute a new grammar

We intersect our automaton F with our grammar G

The resulting grammar G' also generates all REs for the User Resolved Object, with one extra property:

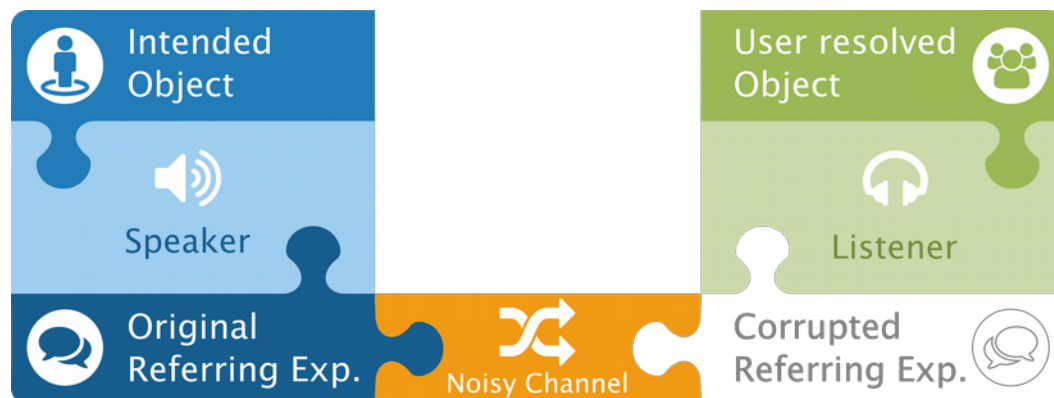
The cost of generating r from G' equals the edit distance between r and our original RE

Step 3: Compute a new grammar

Using Viterbi, we obtain the minimal-distance RE of G'

This is the RE that refers to the User Resolved Object and has the minimum edit cost to our RE

And this is how we find both our missing RE, and the corruption from the Noisy Channel



Example



Original RE	the	yellow	cup	to the right of	the		jar
Edit sequence	<i>Keep</i>	<i>Delete</i>	<i>Keep</i>	<i>Subst.</i>	<i>Keep</i>	<i>Insert</i>	<i>Subst.</i>
Corrupted RE	the		cup	to the left of	the	water	boiler

Example



Original RE	the	yellow	cup	to the right of	the	jar
Edit sequence	<i>Keep</i>	<i>Subst.</i>	<i>Keep</i>	<i>Keep</i>	<i>Keep</i>	<i>Keep</i>
Corrupted RE	the	white	cup	to the right of	the	jar

Example



The cup to the right of the yellow cup

The cup to the right of the cup

The Berlin cup

The white cup

The cup to the left of the water boiler

The cup above the folder with
a red label

The cup in front of the green book

The cup in front of the cup

The white cup to the right of the jar

The cup without squares

The cup to the right of the cup to the
right of the jar

...



PART II

Strategies for contrastive feedback

Feedback strategies

Emphasis

Add contrastive focus to words which were changed by the corruption

Original RE	the	yellow	cup	to the right of	the	jar
Edit sequence	<i>Keep</i>	<i>Subst.</i>	<i>Keep</i>	<i>Keep</i>	<i>Keep</i>	<i>Keep</i>
Corrupted RE	the	white	cup	to the right of	the	jar

*No, the **YELLOW** cup to the right of the jar*

Feedback strategies

Shortening

Remove constituents which were not changed by the corruption

Original RE	the	yellow	cup	to the right of	the	jar
Edit sequence	<i>Keep</i>	<i>Subst.</i>	<i>Keep</i>	<i>Keep</i>	<i>Keep</i>	<i>Keep</i>
Corrupted RE	the	white	cup	to the right of	the	jar

*No, the **YELLOW** cup*

Baseline strategies

Repeat

Present the same Original RE

Original RE	the	yellow	cup	to the right of	the	jar
Edit sequence	<i>Keep</i>	<i>Subst.</i>	<i>Keep</i>	<i>Keep</i>	<i>Keep</i>	<i>Keep</i>
Corrupted RE	the	white	cup	to the right of	the	jar

No, the yellow cup to the right of the jar

Baseline strategies

Random

Add contrastive focus to random adjectives, adverbs, and/or prepositions

Original RE	the	yellow	cup	to the right of	the	jar
Edit sequence	<i>Keep</i>	<i>Subst.</i>	<i>Keep</i>	<i>Keep</i>	<i>Keep</i>	<i>Keep</i>
Corrupted RE	the	white	cup	to the right of	the	jar

*No, the yellow cup to the **RIGHT** of the jar*



PART III

Experiments and results

Experimental setup



GIVE Challenge

Koller et al. Report on the Second NLG Challenge on Generating Instructions in Virtual Environments. (INLG 2010)



TUNA People corpus

van der Sluis, Gatt, and van Deemter. Evaluating algorithms for the generation of referring expressions: Going beyond toy domains. (RANLP 2007)



Experimental setup

Crowdsourced overhearer experiment

Players were asked to rate which one
of two REs was *better*



We wanted our player to
select this button.
So we told them...

*Press the button
to the left of the chair*



But they selected this button
instead.

Which correction is
better for this scene?

- a. No, press the button to the left of the chair*
- b. No, press the button to the **LEFT** of the chair*



We wanted our player to select
the person circled in green.
So we told them...

*The dark haired bearded
young man in a suit
wearing a tie*

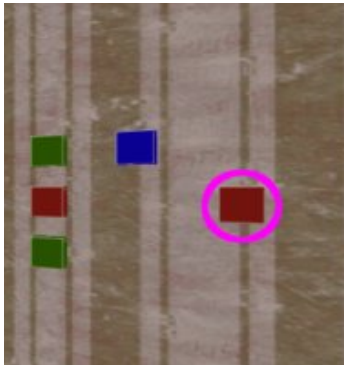
But they selected the person
circled in red instead.
Which correction is better for
this scene?

- a. No, the man **IN A SUIT WEARING A TIE**
- b. No, the dark haired **BEARDED** young man in a suit
wearing a tie

Experimental setup

Crowdsourced overhearer experiment

Players were asked to rate which one
of two REs was *better*

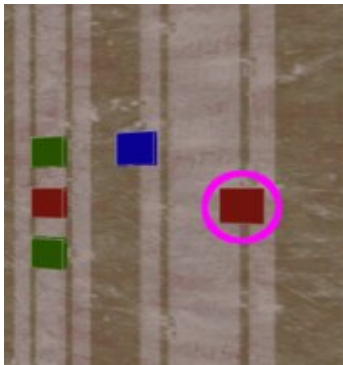


16 scenes
142 subjects
943 judgements



8 scenes
65 subjects
240 judgements

Results – GIVE Challenge



	Repeat	Random	Shortening
Random	-0.041	–	–
Shortening	0.141	0.109	–
Emphasis	0.570***	0.600***	0.376***

Emphasis \geq Shortening

Highly significant

Emphasis \geq Repeat

Highly significant

Emphasis \geq Random

Highly significant

Shortening \geq Repeat

Not significant

Shortening \geq Random

Not significant

Repeat \geq Random

Not significant

Formula: $\frac{(\text{Strat}_1 \text{ pref. over Strat}_2) - (\text{Strat}_2 \text{ pref. over Strat}_1)}{(\text{tests between Strat}_1 \text{ and Strat}_2)}$

Significance at $p < 0.001$,
two-tailed binomial test over preference counts

Results – TUNA People Corpus



	Repeat	Random
Random	0.425***	–
Emphasis	0.575***	0.425***

Emphasis \geq Random

Highly significant

Emphasis \geq Repeat

Highly significant

Random \geq Repeat

Highly significant

Formula: $\frac{(\text{Strat}_1 \text{ pref. over Strat}_2) - (\text{Strat}_2 \text{ pref. over Strat}_1)}{(\text{tests between Strat}_1 \text{ and Strat}_2)}$

Significance at $p < 0.001$,
two-tailed binomial test over preference counts



Conclusion

We generate contrastive REs in an effective and efficient way

Results show a clear preference for our main strategy over all baselines

Our approach can be easily adapted to many RE generation system



Thank you for your attention

Martín Villalba
villalba@coli.uni-saarland.de



Saarland University

