

Same Author or Just Same Topic?

Towards Content-Independent Style Representations

Anna Wegmann, Marijn Schraagen & Dong Nguyen
Department of Information and Computing Sciences
Utrecht University, Utrecht, the Netherlands

June 15th 2022, SCALE 2022



Utrecht
University

Content-Independent Style Representations



Try me!

⚡ Hosted inference API ⓘ

📄 Sentence Similarity

Examples



Source Sentence

r u a fan of them or something?

Sentences to compare to

Are you one of their fans?

Oh yea and tha young dr got a bad haircut.

Style in Natural Language Processing

Style is an integral
part of language

- influences perception

e.g., [El Baff et al., Analyzing the Effect of Style in News Editorial Argumentation, 2020]

- relevant for NLU & NLG

e.g., [Danescu-Niculescu-Mizil et al., Chameleons in Imagined Conversations: A New Approach to Understanding Coordination of Linguistic Style in Dialogs, 2011], [Ficler and Goldberg, Controlling Linguistic Style Aspects in Neural Language Generation, 2017]

Style in Natural Language Processing

Style is an integral
part of language

- influences perception

e.g., [El Baff et al., Analyzing the Effect of Style in News Editorial Argumentation, 2020]

- relevant for NLU & NLG

e.g., [Danescu-Niculescu-Mizil et al., Chameleons in Imagined Conversations: A New Approach to Understanding Coordination of Linguistic Style in Dialogs, 2011], [Ficler and Goldberg, Controlling Linguistic Style Aspects in Neural Language Generation, 2017]

Computational representations
focus on semantics & syntax

Learning Style Representations

Authorship Verification (AV)

as training task:

Learning Style Representations

Authorship Verification (AV)

as training task:

True. Some of them
are religious nutbars.

Lol all i can say is got-
tem. you got wrecked
by a class A douchebag

I mean Judaism and
Sikhi also believe that.

Label: ■ Same Author

Learning Style Representations

Authorship Verification (AV)

+ extensive training

e.g., [Zhu et al., Idiosyncratic but not arbitrary: Learning idiolects in online registers reveals distinctive yet consistent individual styles., 2021; Rivera-Soto et al., Learning Universal Authorship Representations, 2021]

as training task:

True. Some of them
are religious nutbars.

Lol all i can say is got-
tem. you got wrecked
by a class A douchebag

I mean Judaism and
Sikhi also believe that.

Label: ■ Same Author

Domain: ■ r/OkCupid ■ r/DebateReligion

Learning Style Representations

Authorship Verification (AV)

+ extensive training

e.g., [Zhu et al., Idiosyncratic but not arbitrary: Learning idiolects in online registers reveals distinctive yet consistent individual styles., 2021; Rivera-Soto et al., Learning Universal Authorship Representations, 2021]

⚡ content \propto style

e.g., [Sundararajan et al., What represents “style” in authorship attribution?, 2018; Sari et al., Topic or style? Exploring the most useful features for authorship attribution., 2018]

as training task:

True. Some of them are **religious** nut-bars.

Lol all i can say is gottem. you got wrecked by a class A douchebag

I mean **Judaism** and **Sikhi** also **believe** that.

Label: ■ Same Author

Domain: ■ r/OkCupid ■ r/DebateReligion

Learning with Content Control

Authorship Verification (AV)

as training task:

Learning with Content Control

Authorship Verification (AV)

as training task:

don't suggest an open relationship if you're not ready

Aren't open relationships usually just about fixing something in the relationship?

it's clear that these are wildly different situations

Src: ■ conversation C_1 ■ conversation C_2

Learning with Content Control

Authorship Verification (AV)

- using **conversation** as content control

as training task:

don't suggest an open relationship if you're not ready

Aren't open relationships usually just about fixing something in the relationship?

it's clear that these are wildly different situations

Src: ■ conversation C_1 ■ conversation C_2

Learning with Content Control

Authorship Verification (AV)

- using **conversation** as content control
 - on 546,000 Reddit utterances written by 194,000 authors
- [Baumgartner et al., The Pushshift Reddit dataset, 2020]

- using triplet loss to fine-tune RoBERTa with SBERT

[Reimers et al., Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks, 2019]

as training task:

don't suggest an open relationship if you're not ready

Aren't open relationships usually just about fixing something in the relationship?

it's clear that these are wildly different situations

Src: ■ conversation C_1 ■ conversation C_2

Evaluation

1. Authorship Verification

Are same author sentences close to each other independent from domain & conversation?

2. Content Independence

Are known style dimensions (e.g., formal vs. informal) preferred over content information?

Selected Results — Authorship Verification

- ▶ Conversation content control tasks are hardest to solve
- ▶ Decreasing standard deviation with content control

Training Task	No	Do AUC	Co
RoBERTa base	0.61	0.57	0.53
No Content Control	0.77	0.63	0.58
Domain (Do)	0.72	0.70	0.68
Conversation (Co)	0.71	0.70	0.69

Content Independence - STEL-Or-Content Task

r u a fan of them
or something?

Are you one of their
fans?

Oh yea and that young
dr got bad hair

Same Style Paraphrase

based on [Wegmann and Nguyen, Does it capture STEL?, EMNLP 2021]

Content Independence - STEL-Or-Content Task

r u a fan of them
or something?

Are you one of their
fans?

Oh yea and that young
dr got bad hair

Same Style Paraphrase

based on [Wegmann and Nguyen, Does it capture STEL?, EMNLP 2021]

- formal/informal

Yahoo! Answers [Rao et al., Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer, 2018]

- (no) numb3r substitution

based on **Reddit** utterances

- complex/simple

Wikipedia [Xu et al., Optimizing Statistical Machine Translation for Text Simplification, 2016]

- (no) con'raction

based on **Wikipedia** abstracts

Selected Results — Content Independence

**Conversation
content control
leads to highest
content
independence**

Training Task	STEL- Or-Content Task (ACC)
RoBERTa base	.05
No Content Control	.24
Domain (Slide 3)	.32
Conversation (Slide 4)	.42

Summary

- ▶ Usually only limited control for content in Style/AID representations
- ▶ We control for content with same conversation variable
- ▶ Resulting in more content-independent representations

Summary

- ▶ Usually only limited control for content in Style/AID representations
- ▶ We control for content with same conversation variable
- ▶ Resulting in more content-independent representations

⇒ **potential to increase classification performances for diverse demographics & social groups**

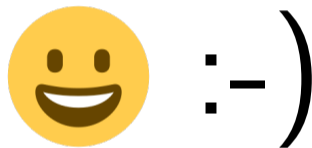
Same Author or Just Same Topic? @SCALE22



a.m.wegmann@uu.nl

@anna_wegmann

- **Evaluating representations:** extending STEL, e.g., with AID tasks; standardized AID tasks in different conditions



Same Author or Just Same Topic? @SCALE22



a.m.wegmann@uu.nl

@anna_wegmann

- **Evaluating representations:** extending STEL, e.g., with AID tasks; standardized AID tasks in different conditions
- **Interpreting representations:** clustering, probing

C #	Consistent	Example
3	no last punct.	I am living in china, they are experiencing an enormous baby boom
4	punctuation / casing	huh thats odd i'm in the 97% percentile on iq tests, the sat, and the act
5	' vs '	I assume it's the blind lady?
7	linebreaks	I admire what you're doing but [...] I know I'm [...]

Same Author or Just Same Topic? @SCALE22



a.m.wegmann@uu.nl

@anna_wegmann

- **Evaluating representations:** extending STEL, e.g., with AID tasks; standardized AID tasks in different conditions
- **Interpreting representations:** clustering, probing
- **Training topic-independent models:** hard(er) training examples, e.g., with high semantic/style similarity from previous models

Same Author or Just Same Topic? @SCALE22



a.m.wegmann@uu.nl
@anna_wegmann

- **Evaluating representations:** extending STEL, e.g., with AID tasks; standardized AID tasks in different conditions
- **Interpreting representations:** clustering, probing
- **Training topic-independent models:** hard(er) training examples, e.g., with high semantic/style similarity from previous models
- **AID and Style:** generalization of AID to linguistic style representations? Idiolect, audience design, linguistic accommodation, ...

Data & Code: github.com/nlpsoc/style-embeddings

Task Statistics – 600 conversations from 100 subreddits

CC level	Data Split	Setup		Utterance #	Author		(A_1, A_2)		(A_1, B)	
		# AV	# CAV		#	ma	co	do	co	do
Conversation	train set	420,000	210,000	546,757	194,836	9	0.27	0.56	1.00	1.00
	dev set	90,000	45,000	116,451	41,848	8	0.26	0.55	1.00	1.00
	test set	90,000	45,000	116,621	41,902	8	0.27	0.55	1.00	1.00
Domain	train set	420,000	210,000	544,587	240,065	9	same pairs	0.01	1.00	
	dev set	90,000	45,000	116,490	50,939	8	as	0.02	1.00	
	test set	90,000	45,000	116,586	51,182	8	conversation	0.02	1.00	
No	train set	420,000	210,000	548,082	270,079	9	same pairs	0.00	0.01	
	dev set	90,000	45,000	117,149	57,352	8	as	0.00	0.01	
	test set	90,000	45,000	117,434	57,726	8	conversation	0.00	0.02	

Table 1: **Data Split Statistics.** Per content control (CC) level, we display the number of tasks per setup (# CAV, # AV), unique utterances and authors for each split. We also show the maximum number of times an author occurs as A_1 's author (ma) and the fraction of same author (A_1, A_2) and utterance pairs of different authors (A_1, B) that occur in the same conversation (co) and domain (do).

STEL Task Generation*

*Examples shortened

formal	informal
Are you one of their fans?	<u>r u</u> a fan of them or something?
Oh, and also that young physician got an unflattering haircut	Oh <u>yea</u> and that young <u>dr</u> got a bad haircut

- formal/informal

Yahoo! Answers paraphrases from [Rao and Tetreault, Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer, 2018]

STEL Task Generation*

*Examples shortened

formal	informal
Are you one of their fans?	<u>r u</u> a fan of them or something?
Oh, and also that young physician got an unflattering haircut	Oh <u>yea</u> and that young <u>dr</u> got a bad haircut



r u a fan of them or something?

Oh, and also that young physician got an unflattering haircut

Are you one of their fans?

Oh yea and that young dr got a bad haircut

- formal/informal

Yahoo! Answers paraphrases from [Rao and Tetreault, Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer, 2018]

STEL Task Generation*

*Examples shortened

formal	informal
Are you one of their fans?	<u>r u</u> a fan of them or something?
Oh, and also that young physician got an unflattering haircut	Oh <u>yea</u> and that young <u>dr</u> got a bad haircut



r u a fan of them or something?

Are you one of their fans?

Oh, and also that young physician got an unflattering haircut

Oh yea and that young dr got a bad haircut

- formal/informal

Yahoo! Answers paraphrases from [Rao and Tetreault, Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer, 2018]

STEL Task Generation*

*Examples shortened

simple	complex
These rock formations are made of sandstone.	These rock formations are <u>composed</u> of sandstone.
The Odyssey is an old poem written by Homer.	The Odyssey is an <u>ancient</u> poem <u>attributed</u> to Homer.



These rock formations are composed of sandstone.

These rock formations are made of sandstone.

The Odyssey is an ancient poem attributed to Homer.

The Odyssey is an old poem written by Homer.



- formal/informal

Yahoo! Answers paraphrases from [Rao and Tetreault, Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer, 2018]

- complex/simple

Wikipedia paraphrases from [Xu et al., Optimizing Statistical Machine Translation for Text Simplification, 2016]

STEL Task Generation*

*Examples shortened

no nb3r	nb3r
<3 friends for-ever	<3 friends <u>4</u> ever
Dude 30\$ is heaps cheap	<u>D00d</u> 30\$ is heaps cheap



<3 friends forever

<3 friends 4ever

D00d 30\$ is heaps cheap

Dude 30\$ is heaps cheap

- formal/informal

Yahoo! Answers paraphrases from [Rao and Tetreault, Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer, 2018]

- **(no) numb3r substitution**

based on **Reddit** utterances

- complex/simple

Wikipedia paraphrases from [Xu et al., Optimizing Statistical Machine Translation for Text Simplification, 2016]

STEL Task Generation*

*Examples shortened

no con'raction	con'raction
It has become one of the world's most significant cities.	<u>It's</u> become one of the world's most significant cities.
Will does not refer to any particular desire.	Will <u>doesn't</u> refer to any particular desire.



It has become one of the world's most significant cities.

It's become one of the world's most significant cities.

Will does not refer to any particular desire.

Will doesn't refer to any particular desire.



- formal/informal

Yahoo! Answers paraphrases from [Rao and Tetreault, Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer, 2018]

- (no) numb3r substitution

based on **Reddit** utterances

- complex/simple

Wikipedia paraphrases from [Xu et al., Optimizing Statistical Machine Translation for Text Simplification, 2016]

- (no) con'raction

based on **Wikipedia** abstracts

STEL Task Generation*

*Examples shortened

Style 1	Style 2
Anchor Sentence 1	Anchor Sentence 2
Alternative Sentence 2	Alternative Sentence 1

Paraphrases



Anchor Sentence 1

Anchor Sentence 2

Alternative Sentence 1

Alternative Sentence 2 ✓

- formal/informal

Yahoo! Answers paraphrases from [Rao and Tetreault, Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer, 2018]

- (no) number substitution

based on **Reddit** utterances

- complex/simple

Wikipedia paraphrases from [Xu et al., Optimizing Statistical Machine Translation for Text Simplification, 2016]

- (no) contraction

based on **Wikipedia** abstracts

STEL Task Generation*

*Examples shortened

Style 1	Style 2
Anchor Sentence 1	Anchor Sentence 2
Alternative Sentence 2	Alternative Sentence 1

Paraphrases



Anchor Sentence 1

Anchor Sentence 2

Alternative Sentence 1

Alternative Sentence 2 ✓

STEL is modular:

- formal/informal

Yahoo! Answers paraphrases from [Rao and Tetreault, Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer, 2018]

- (no) number substitution
based on **Reddit** utterances

- complex/simple

Wikipedia paraphrases from [Xu et al., Optimizing Statistical Machine Translation for Text Simplification, 2016]

- (no) contraction
based on **Wikipedia** abstracts

STEL Task Generation*

*Examples shortened

Style 1	Style 2
Anchor Sentence 1	Anchor Sentence 2
Alternative Sentence 2	Alternative Sentence 1

Paraphrases



Anchor Sentence 1

Anchor Sentence 2

Alternative Sentence 1

Alternative Sentence 2



STEL is modular:

complex { - formal/informal
- **Yahoo! Answers** paraphrases from [Rao and Tetreault, Dear Sir or Madam, May I Introduce the GYAFC Dataset: Corpus, Benchmarks and Metrics for Formality Style Transfer, 2018]

simple { - (no) num3r substitution
- based on **Reddit** utterances

- complex/simple
- **Wikipedia** paraphrases from [Xu et al., Optimizing Statistical Machine Translation for Text Simplification, 2016]

- (no) con'raction
- based on **Wikipedia** abstracts