Comparing British and American Audio Description of Movies

Igor Sterner, Alex Lascarides, Frank Keller

THE UNIVERSITY of EDINBURGH
informatics

Narrative in Movies

Here is a scene in *Oppenheimer (2023)* with Audio Description

https://www.dropbox.com/scl/fi/hj4ze0h3ly16sikeo042z/oppenheimer-apple-subs.mp4?rlkey=z61pkqr2awfh267kgxw0tiffk&st=229bjfdb&dl=0

Audio Description (AD) as a narrative art form

- AD is a verbal narration of key visual content necessary to follow the story
- It provides access to blind, visually impaired and other audiences

AD is a translation of a visual narrative into a verbal one

Audio Description (AD) as a narrative art form

- AD is a verbal narration of key visual content necessary to follow the story
- It provides access to blind, visually impaired and other audiences

> AD is a translation of a visual narrative into a verbal one

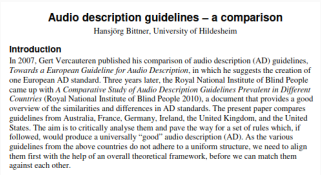
- AD is a narrative and literary art form
- The degree of creative license is hotly debated

Why AD?

- AD makes the underlying processes of narrative construction in movies explicit
- We aim to understand and model movie narratives
- ...with application to the automatic generation of AD

Comparing British and American AD of movies

- Prior comparisons of British and American AD rely on theory and anecdote
- We provide the first quantitative comparison of the two regions



Audio description guidelines – a comparison

Hansjörg Bittner, University of Hildesheim

Introduction

In 2007, Gert Vercauteren published his comparison of audio description (AD) guidelines, *Towards a European Guideline for Audio Description*, in which he suggests the creation of one European AD standard. Three years later, the Royal National Institute of Blind People came up with *A Comparative Study of Audio Description Guidelines Prevalent in Different Countries* (Royal National Institute of Blind People 2010), a document that provides a good overview of the similarities and differences in AD standards. The present paper compares guidelines from Australia, France, Germany, Ireland, the United Kingdom, and the United States. The aim is to critically analyse them and pave the way for a set of rules which, if followed, would produce a universally "good" audio description (AD). As the various guidelines from the above countries do not adhere to a uniform structure, we need to align them first with the help of an overall theoretical framework, before we can match them against each other.

RED BEE

About Us Services

News 05 May, 2025

Crossing The Pond: A Deep Dive Into UK And US Audio Description

1. We define hypotheses based on regional guidelines and the stated experiences of AD practitioners,
2. design automatic tests,
3. and apply them to a controlled corpus

Hypothesis: scene changes and lexical preferences

Scene changes

Inside a tent, Lawrence drinks from a bottle and passes it to Frank. [UK]

After nightfall, [...] Lawrence passes a liquor bottle [...] [US]

In a tent at night, [...] Lawrence passes a bottle of booze [...] [US→UK post-edit]

Hypothesis: scene changes and lexical preferences

Lexical preferences

Inside a tent, Lawrence drinks from a **bottle** and passes it to Frank. [UK]

After nightfall, [...] Lawrence passes a **liquor bottle** [...] [US]

In a tent at night, [...] Lawrence passes a **bottle** of booze [...] [US→UK post-edit]

Hypothesis: voice and aspect

Voice

Oppenheimer sees a torch **being shone** at cars parked on a suburban street. [UK]
Oppenheimer notices someone with a flashlight checking out a car [...]. [US]
Oppenheimer notices someone shining a torch by a car [...]. [US→UK post-edit]

Aspect

Oppenheimer **is standing** at the lectern. [UK]
Oppenheimer **stands** at a lectern. [US]
Oppenheimer **stands** at a lectern looking slightly nervous. [US→UK post-edit]

Hypothesis: subjectivity

Subjectivity

Oppenheimer is standing at the lectern. [UK]

Oppenheimer stands at a lectern. [US]

Oppenheimer stands at a lectern **looking slightly nervous**. [US→UK post-edit]

Naming in advance

A smile from **Evans**, the oldest board member. [UK]

An **elderly member of the board** smiles. [+8 mins] Robb shakes hands with the chairman, the elderly **Evans** and [...]. [US]

An older member of the board, **Ward Evans**, smiles. [US→UK post-edit]

Hypothesis: dialogue overlap

The green apple from earlier

As classmates leave the lab, one shouts to Oppenheimer, *"Don't forget to clean up!"*

- **US AD** voices over this dialogue entirely to describe a green apple
- **UK AD** does not overlap

The REFRAMED dataset

(note that this dataset is presented in another paper, which is currently in submission)

- Professional American AD, British AD and dialogue subtitles for 206 movies
- (also: 3K+ scenes from the movies, aligned screenplays, and SDH)
- Fully public release for scene-based data; academic release for full-movie data

	British AD	American AD
Total token count	1,137,598	1,419,576
Total sentence count	121,123	171,224
Average sentences per movie	588	831
Average tokens per movie	5,522	6,891

The REFRAMED dataset

(note that this dataset is presented in another paper, which is currently in submission)

- Professional American AD, British AD and dialogue subtitles for 206 movies
- (also: 3K+ scenes from the movies, aligned screenplays, and SDH)
- Fully public release for scene-based data; academic release for full-movie data

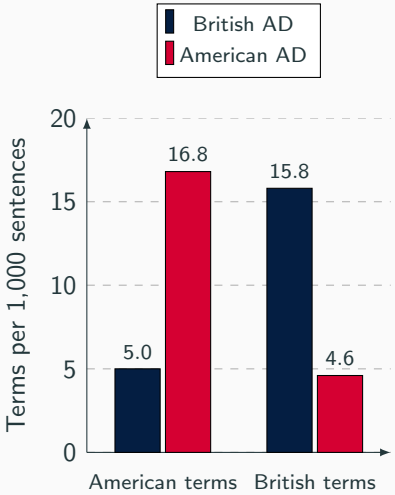
	British AD	American AD
Total token count	1,137,598	**1,419,576**
Total sentence count	121,123	**171,224**
Average sentences per movie	588	**831**
Average tokens per movie	5,522	**6,891**

Automatic analysis

- Parse every AD and dialogue sentence with Stanza: (tokenization, sentence segmentation, universal dependencies and POS tags, OntoNotes named entities)
- Match against pre-defined lists of American and British terms
- One automatic test per hypothesis, e.g. present progressive via a present-tense auxiliary with a participle; passive voice via passive dependency relations
- All tests are relatively simple; see paper for more details

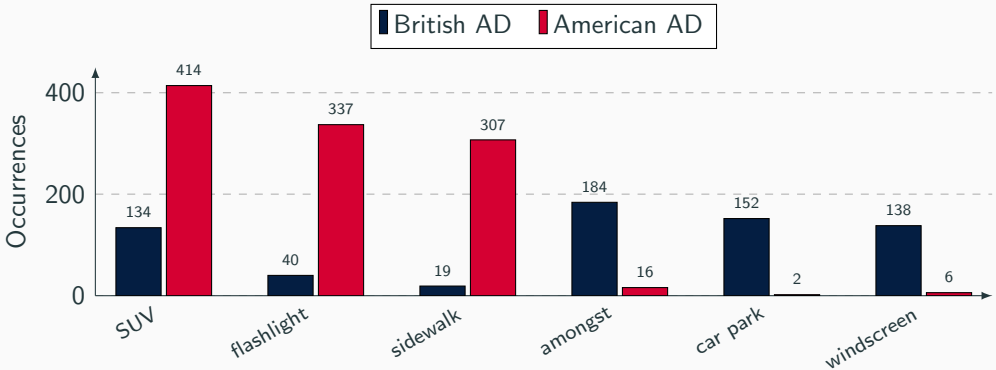
Results: lexicon

(all differences in upcoming results are statistically significant)

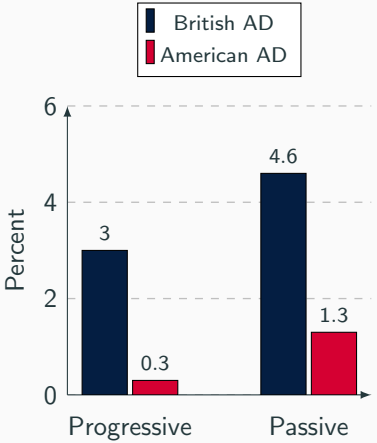


- Each region uses its own terms about three times more frequently
- Differences only relevant for c. 1% of sentences

Results: lexical examples

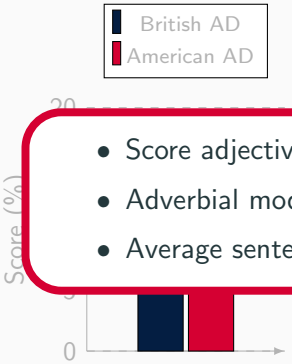


Results: aspect and voice



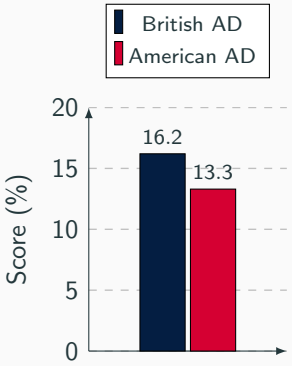
- British AD uses the progressive aspect 10x
- Passive voice is also more frequent in British AD
- US guidelines are more explicit than UK in using *simple* language

Results: subjectivity



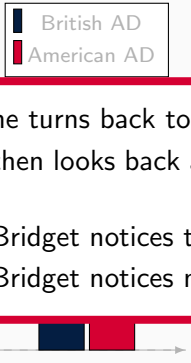
- Score adjectives using a lexicon of subjectivity ratings
- Adverbial modifiers also contribute to the sentence score
- Average sentence-level scores within each movie

Results: subjectivity



- British AD has higher average adjective subjectivity.
- *beautiful* occurs almost three times as often in British AD
- US guidelines directly address the use of this word, negatively stating that it is vague and evaluative

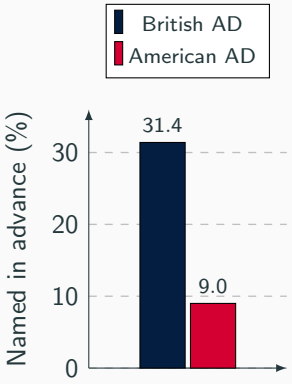
Results: subjectivity



he turns back to find a **beautiful** woman glaring at him. [UK]
then looks back and finds a **dark-haired young** woman standing before him. [US]

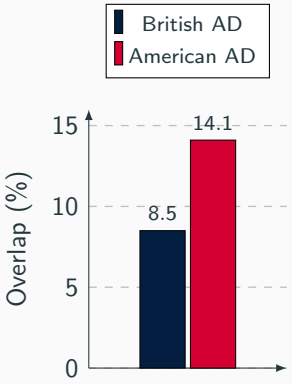
Bridget notices the church is mainly filled with **beautiful** young women. [UK]
Bridget notices many of the pews filled with **fashionable** young women. [US]

Results: naming in advance



- British AD names characters in advance more frequently
- In British AD, the average named-in-advance character is 11 mins earlier, vs. 7 mins in US AD
- Our analysis shows that it is still used in US AD

Results: dialogue overlap



- American AD overlaps subtitles and music lyrics more often
- Then again, it is also longer overall

Results: cueing scene changes

We train a model to detect scene boundaries from AD text alone, then test on each region. Recall scores:

		Test	
		UK	US
Train	UK	55.4	**72.4**
	US	42.9	**66.0**
	UK+US	55.1	**69.0**

(fixed number of train samples)

- Recall is significantly lower on UK AD
- The model learns explicit cues that are more frequent in American AD
- British AD verbalises scene changes less explicitly

Take-home messages

- British and American AD show small but statistically distinguishable differences
- Every difference is motivated by the regional guidelines (except lexical variation)
- AD is a constrained narrative technique
- It makes narrative construction in movies explicit: attractive for narratology

- This work is part of my PhD on AD generation! In the coming months, I will release the REFRAMED dataset – please do speak to me further about this