[Training]

Query: House **drink** his **coffee**

Query: Ryan takes a **drink** of his **coffee**

Video Corpus

Training data

12.33 sec

13.77 sec — 14.79 sec

Ground-truth

Biased annotations

Query: Beckett pours a **coffee** into her cup

"coffee !"

(Retrieval Bias)

Scene of drinking coffee

[Inference]

Query: Beckett dumps coffee from cup into garbage can.

Video Corpus

2.73 sec — 4.1 sec   6.0 sec — 9.0 sec

Ground-truth   **Prediction**

"Scene of dumping coffee"   "Scene of drinking coffee"

(Retrieval Bias)