

概述

本文件为标注员提供了一个全面的指南，帮助准确识别和分类讽刺(sarcasm)、真诚表达(sincere)和非讽刺(non-sarcasm)的语句。讽刺作为一种非字面语言形式，其特征在于说话者表面言辞与潜在意图之间的差异，通常带有负面情绪并针对一个目标。文件中的示例展示了说话者通过不同方式表达讽刺的手法，包括文字、语气和面部表情等。以下内容具体阐释了讽刺的定义、例子和非讽刺实例，并提供了测试案例供标注参考。

第一部分：讽刺的定义与特征

- 讽刺的基本定义

Sarcasm are expressions that convey meanings distinct from the literal statement, often used for self-deprecation, critique, mocking, or humor. Sarcasm can be expressed unimodally or multimodally through words, tone, facial expressions, and body languages. Sarcastic remarks can also be conveyed through rhetorical devices like hyperbole, understatement and mockery imitation. Our guideline provides annotators comprehensive key features of sarcasm while allowing the flexibility to interpret. To maintain consistency, we held weekly meetings to review examples and address ambiguous cases, progressively refining alignment.

讽刺是通过表达与字面意思不同的含义来传递信息，通常用于自嘲、批评、讽刺或幽默。讽刺可以通过单一或多模态(如语言、语调、面部表情和肢体语言)来表达。讽刺的言论也可以通过修辞手法，如夸张、轻描淡写和模仿讽刺来传达。比如，假设请求被朋友拒绝了之后，说话人用“你真是个乐于助人啊”来表达对对方的不满情绪(潜在意图)，而不是褒奖(“乐于助人”的表面意义)。

- 讽刺的必要特征 —

- 表面言辞与潜在意图的差异
- 表达个人评价(目的：嘲笑，批评，自嘲，幽默)
- 可以透过字面，语气，或是面部表情等表达。
- 可使用修辞手法来表达：重复，模仿，比喻等

第二部分：讽刺的表达方式

说话人通过不同的方式来传达讽刺：

第一是通过字面，比如下面的Example1，“天使轮”是褒义词，而用来搭配“出事儿”这个含有负面意义的词，这使得听者了解到“天使轮”并不是使用其表面的正面意义，而说话人的潜在意图是与之有差异甚至相反是贬义的——“天使轮”是讽刺的用法。另外一个例子是“太棒了，我真的特别需要这种倒霉事。”其中“太棒了”并没有发挥其正面的表面意义，通过与“倒霉事”的搭配，听者了解到这

是讽刺的用法。其他的例子也包括使用讽刺常见的短语或者词，例如，“真是一点儿都不麻烦！”这类句子即使在没有具体语境的情况下，通常也能直接被理解为讽刺。

第二是通过语气，例如下面的Example3，“我们家是非常喜欢女儿的”本来可以按照字面的意思理解，但是说话人放缓了速度，并且用了很平淡的语气，让听者了解到这里并没有使用字面意义，而潜在意思是与之不同甚至相反的---说话人用了讽刺。

第三是通过面部表情，例如Example2，说话人在说出“但是我现在混的一般，一无所有，有一点羡慕那些留在家乡的姐妹们了啊，她们连前夫都有了。”这句话时，用了抿嘴和放空的表情，让听者了解到潜在意思与字面意思有差异---讽刺。

Example1: [\[link\]](#)

Timestamps: 13:07-13:26

Transcripts: 我到剧场报名演出，老板只看了我一眼，当场签约，我刚签完字，他就让助理给我买500万的人身意外保险，这个老板太有经济头脑了，只要是出点啥事儿，我就是天使轮。

Label: Sarcasm

Source of sarcasm: text

Intensity: 1

Example2: [\[link\]](#)

Timestamps: 0:14 - 0:32

Transcripts: 真的，上电视，对我来说，太难了，太难了，因为我也是一个农村人。我真的是一个地地道道的来自于大山的一个女孩儿，但是我是我家族里面第一个考上大学的人。但是我现在混的一般，一无所有，有一点羡慕那些留在家乡的姐妹们了啊，她们连前夫都有了。

Label: sarcasm

Source of sarcasm: visual

Intensity: 2

Example3: [\[link\]](#)

Timestamps: 0:45 - 0:53

Transcripts: 我们家有四个孩子，大姐，二姐，我，还有我弟弟，一看这样的排列组合你们就知道，我们家是非常喜欢女儿的。

Label: sarcasm

Source of sarcasm: visual, audio

Intensity: 3

第三部分：非讽刺的例子

而什么又不属于讽刺呢？

Example 6中，说话人通过夸张模仿参加综艺的人，来表达对他们行为的嘲笑。虽然这与讽刺的一致（嘲笑或者中伤语境中的victim），但是并没有使用讽刺的手法--使用中性或者褒义的词或者句子来表达负面的潜在意思。所以这也不是讽刺。

Example 6 [\[link\]](#)

Timestamps: 55:39-55:58

Transcripts: 我觉得，啊对不起，有时候，相爱和分开不一定是相悖的，我说，哈？我都听不懂。

第四部分：真诚表达的例子

另外，为了建立一个平衡的数据库，我们也需要一些真实表达意图的句子(sincere)，以下是一些例子。这类型的句子就是说话人在表达自己真诚真实的意图。

Example 7 [\[link\]](#)

Timestamps: 1:01:01-1:01:12

Transcripts: 现在有时候看社交媒体，我看有很多传递身材焦虑的言论。比如说有很多穿搭博主，我觉得他们特别擅长给身上的肉起名字但是又非常不好理解。

Label: Sincere

Example 8 [\[link\]](#)

Timestamps: 14:15-14:33

Transcripts: 而且我谈的也挺黄昏恋的，就是散布，钓鱼，逛公园，非常平淡，我跟你们说，我现在就喜欢这种淡爱，因为那种轰轰烈烈的爱情，我已经从我闺蜜那里参与过很多次了。

Label: Sincere

第五部分：测试案例

以上就是关于讽刺的含义的具体解释，现在可以尝试标注以下这几个句子，你觉得他们是sarcasm还是sincere或者non sarcasm？

Test 1 [\[link\]](#)

Timestamps: 54:27-54:59

Transcripts: 我真的，我越想越心动，我仔细去看了一下那个离婚综艺，我看完以后很难受，发现我根本就不配上。怎么大家都是两口子，他们的矛盾，就那么体面呢？一对对夫妻，从不探讨房贷谁还孩子谁带，天天不是看海就是野餐，离个婚，比我结婚都浪漫。

Label:

Test 2 [\[link\]](#)

Timestamps: 3:50-4:04

Transcripts: 你看那个奥运会，其实不需要那么复杂的赛制，跑就是了，跳就是了，奥运会跟我们这个比赛一比，你就会发现，高端的赛事往往只需要最朴素的赛制，百米不会搞什么，大快圈，微快，圈，好快圈。

Label:

Test 3 [\[link\]](#)

Timestamps: 5:36-5:52

Transcripts: 你想想看其他公司，什么配领导喝酒，和开心了说可能提拔你，或者给你点业务，脱口秀行业，你陪领导喝开心了，他只会天天叫你陪他喝酒。他不光不给你业务，他耽误你业务。

Label:

Test 4 [\[link\]](#)

Timestamps: 4:21 - 0:53

Transcripts: 我去那个公司，第一次团建，就是给老板他爸过生日，我一分钱还没挣呢，先随了二百，但人家还挺体贴，知道我没钱，说从工资里边直接扣。老板招我这么个员工，真是给家里面谋福利了。

Label: