

Game-Changing Event Definition and Detection in an eSports Corpus

Emily Grace Olshefski
Montclair State University
1 Normal Avenue
Montclair, NJ 07043, USA
olshefskiel@montclair.edu

Abstract

Despite the growing cultural presence of eSports, no corpus contains this genre of entertainment. This paper presents how a preliminary corpus was created from broadcast speech from a professional game of the eSport *Counter-Strike: Global Offensive (CS:GO)*. The corpus was initially annotated following the Automatic Content Extraction (ACE) event subtype definitions for game-changing events: deaths, injuries, and attacks. Event subtype definitions were modified for further annotation to detect a wider range of game-changing events otherwise not defined by ACE. A high degree of inter-annotator agreement for most event subtypes suggests that modifying event subtype definitions for an eSports corpus is necessary to detect the breadth of game-changing events.

1 Introduction

The phenomenon of eSports (electronic sports, or competitive video gaming) is relatively new but has quickly become a global sensation. The advent of local area networks (LAN) has made eSports as competitive, if not more, than traditional sports, especially in countries like South Korea and China and, “this emerging market segment produces billions of dollars and contributes economically to the growth of the sport industry as a whole” (Lee and Schoenstedt, 2011).

Despite the popularity, growth, and cultural impact of eSports, few studies examine the nature of eSports. Furthermore, no linguistic studies of eSports exist due to the lack of eSports as a genre in

corpora. For example, the Manually Annotated Sub-Corpus (American National Corpus Project) contains multiple genres of entertainment, like fiction, essays, and movie scripts. eSports, like film, is a multi-billion dollar industry that is growing rapidly, and by collecting data for eSports other corpora can become representative of emerging language use. While projects such as FrameNet (International Computer Science Institute) and other genre specific corpora like GENIA (Kim et al., 2005), a bi-textmining corpus, have undergone extensive event annotation, eSports has not been included in prior event annotation work, and doing so may provide insight into event detection and definition.

This paper aims to explain how a preliminary eSports corpus was created from the speech of eSports broadcasters who provide play-by-play and color commentary of one of the most popular eSports, *Counter-Strike: Global Offensive*, or *CS:GO*. The corpus was annotated by two annotators with extensive knowledge of not only *CS:GO* but also language the community and broadcasters use to describe the events of *CS:GO* games. Creating and annotating a corpus of *CS:GO* speech also provides a controlled model for the annotation of real-life attack, injury, and death events in, for example, a military based corpus.

The corpus was annotated twice to explore the nature of game-changing events (events that significantly impact the outcome or course of gameplay) in *CS:GO*. The corpus was first annotated following definitions designated by the Automatic Content Extraction (ACE) Program. Annotations were then made after event subtype modification. A new event

subtype was created to more accurately detect game-changing events. This paper is the first step in detecting game-changing events in a corpus comprised entirely of language from eSports.

2 Approach

The first step in the approach to detecting game-changing events in an eSports corpus began with choosing an eSport (*CS:GO*) and creating a corpus from the speech of professional *CS:GO* broadcasters. Annotators were chosen based on knowledge of the eSport, and event definitions and modifications were made to better detect game-changing events.

2.1 *CS:GO* Classic Competitive - Bomb Scenario

Although there are five different game modes in the first-person shooter *CS:GO*, the only one played professionally is Classic Competitive - Bomb Scenario. This game mode is played between a team of five terrorists and five counter-terrorists, loosely modeling a terrorist bomb plant scenario.

The goal of the terrorist team is to plant the bomb and have it explode, and/or to kill all of the counter-terrorists. The goal of the counter-terrorist team is to defuse a planted bomb, kill all of the terrorists, or have a minimum of one player alive in the absence of a bomb plant. The teams switch sides (from counter-terrorist to terrorist, and vice versa) at the halfway mark of 15 rounds, and the first team to win 16 rounds wins the game. In the event of a tie (30 rounds, each team with 15 round wins) the game extends into overtime to determine a winner.

Each player starts each round with 100 health and zero armor points unless they purchase armor in the form of a Kevlar vest and/or helmet. Players lose health and armor points by taking damage from guns, knives, tasers, bomb and grenade explosions, and grenade contact, and damage can be dealt from the opposing team, one's own teammates, or oneself. When a player loses all of their health and armor points they die and are unable to participate in the game until the next round. Professional players often plan their strategies around planting bombs, killing and/or injuring opposing players, and creating space with smoke and flash grenades (referred to as flashbangs in *CS:GO*).

2.2 The *CS:GO* Corpus

The corpus consists of 47 minutes of speech comprised of 10,000 words from an August 2014 video broadcast of a professional *CS:GO* game posted on YouTube between a French team, Titan, and a North American team, Cloud9. The speech was manually transcribed by the author due to a lack of transcripts or closed captioning. This broadcast took place during the Electronic Sports League One Cologne 2014 tournament, and since the broadcasters were not in soundproof booths background noise rendered speech-to-text programs useless in obtaining data.

This video was chosen primarily on the basis of the broadcasters, Auguste "Semmler" Massonant and Anders "Anders" Blume. Semmler and Anders, as they are known in the *CS:GO* community, were chosen due to their expert knowledge of *CS:GO*, their extensive experience broadcasting, and the clarity of their speech. Written permission to use their speech as the basis of this corpus was obtained from both broadcasters. An example of the speech in the corpus is as follows: "Hiko tries to put shots through with the Five Seven, and Seangares looking to do some damage but the bomb will get planted and Cloud9, they just don't have the firepower or the nades to really get in here and have an impact."

2.3 Annotators

The annotation task was completed by two annotators, referred to as Annotator A (the author) and Annotator B. Both annotators have spent over 1000 hours playing *CS:GO*, watch broadcast games weekly, and have not completed any prior annotation tasks.

2.4 Original Event Definitions

Originally the annotators agreed that the ACE English Annotation Guidelines for Events Version 5.4.3 2005.07.01 definitions for LIFE.DIE, LIFE.INJURE, and CONFLICT.ATTACK event subtypes represent the majority of game-changing events in a *CS:GO* game. Player kills (LIFE.DIE events) can significantly alter the outcome of rounds, and even lowering a player's health (LIFE.INJURE events) affects strategy and gameplay. When players are not killing or injuring each other different types

of attack events (CONFLICT.ATTACK) occur that change the course of the game. The ACE (Linguistic Data Consortium, 2005) definitions for these aforementioned event subtypes are as follows:

“An INJURE Event occurs whenever a PERSON Entity experiences physical harm. INJURE Events can be accidental, intentional or self-inflicted.

A DIE Event occurs whenever the life of a PERSON Entity ends. DIE Events can be accidental, intentional or self-inflicted.

An ATTACK Event is defined as a violent physical act causing harm or damage. ATTACK Events include any such Event not covered by the INJURE or DIE subtypes, including Events where there is no stated agent. The ATTACK Event type includes less specific violence-related nouns such as ‘conflict’, ‘clashes’, and ‘fighting’. ‘Gunfire’, which has the qualities of both an Event and a weapon, should always be tagged as an ATTACK Event, if only for the sake of consistency. A ‘coup’ is a kind of ATTACK (and so is a ‘war’).”

2.5 Modified Event Definitions

After the first round of annotation where the annotators followed the ACE definitions, a new event subtype was created in order to better detect more game-changing events and eliminate definition ambiguity. Both annotators agreed that the LIFE.DIE and LIFE.INJURE definitions were unambiguous enough to complete the task.

There was a strong consensus between the annotators, however, that the CONFLICT.ATTACK event subtype did not accurately specify the idea of what an attack always is in *CS:GO*. The annotators took issue with phrasing a CONFLICT.ATTACK event as, “a violent physical act causing harm or damage” when there exist attacks in *CS:GO* that do not cause quantifiable damage or harm but are still violent and physical.

For example, *CS:GO* players have the option of buying grenades, like the HE (high explosive) grenade that explodes on contact, and the Molotov cocktail/incendiary grenade that ignites players and the ground, that are primarily used to injure or kill other players. While these grenades can be used strategically to create space or block access to areas of contention, the general aim of throwing these two types of grenades is to inflict damage. Thus,

Event Subtype	A	B	Agreement
DIE	140	140	0.976
INJURE	1	1	1.00
ATTACK-D	70	63	0.653
ATTACK-ND	22	19	0.952

Table 1: Number of Annotated Events and Degree of Inter-Annotator Agreement

events depicting these types of attack events fit into the ACE definition for CONFLICT.ATTACK.

Smoke grenades, on the other hand, which create a smoke cloud and block vision, and flashbangs, which create a blinding light on the player’s screen if they look in the direction of the grenade as it is thrown, are solely used for a strategic purpose. These attacks do not cause any quantifiable damage but are violent and physical in nature, and despite the ACE definition, would be considered an attack by the *CS:GO* community. The creation of a new event subtype makes a necessary distinction of events that are both game-changing but vary regarding damage and harm. Attacks fitting the old definition were simply relabeled as CONFLICT.ATTACK-D, specifying that the event causes damage. The definition for non-damaging attack events, labeled CONFLICT.ATTACK-ND, is identical to that of CONFLICT.ATTACK-D except the phrase “causing harm or damage” is changed to “cannot or would not cause harm or damage.”

3 Results and Discussion

Table 1 details the total number of game-changing events depicted in the corpus as well as inter-annotator agreement calculated by Cohen’s kappa coefficient. LIFE.DIE events were depicted most in this corpus. Killing opposing players provides a strategic advantage in *CS:GO*, and the majority of games are won by the team with the most kills. The high degree of inter-annotator agreement is due to the familiarity with how the broadcasters generally describe LIFE.DIE events during *CS:GO* games.

Both annotators detected the same LIFE.INJURE event and agreed that although the data for LIFE.INJURE events in this corpus is sparse, the likelihood of encountering more LIFE.INJURE events in a larger corpus is high. Play style varies

greatly in *CS:GO* amongst teams and regions, and including speech from matches between different teams could produce more LIFE.INJURE events.

The lowest degree of agreement between the annotators was for CONFLICT.ATTACK-D events, despite a similar number of events of this subtype detected by both annotators. This lower degree of agreement should be attributed to the fact that certain language was perceived differently, causing one annotator to detect an event where the other did not. For example, in the sentence, “He’s gonna find one headshot, tries to control the spray...” one annotator labeled the act of trying to control gun spray as a CONFLICT.ATTACK-D event and the other annotator did not. However, in other instances where the word ‘spray’ is used as a verb the annotators detected it as a CONFLICT.ATTACK-D event. The annotators agreed that the definition for CONFLICT.ATTACK-D events should be revised to eliminate existing ambiguity, and possibly include a list of words or phrases in *CS:GO* that could signal a CONFLICT.ATTACK-D event.

CONFLICT.ATTACK-ND agreement was also high between annotators despite the relatively small set of event depictions. This high degree of agreement can be attributed to the fact that this event subtype was created specifically to fulfill the need to detect game-changing non-damaging attacks. Regardless, the high degree of inter-annotator agreement indicates that this event subtype definition was generated in such a way that game-changing non-damaging attack events could be consistently and accurately detected. Like LIFE.INJURE both annotators agree that a larger corpus would lead to more CONFLICT.ATTACK-ND events, especially if either team’s strategy relied heavily on the use of non-damaging grenades (flashbangs and smoke grenades).

4 Conclusion and Future Work

The high degree of inter-annotator agreement for the majority of event subtypes demonstrates that modifying event subtype definitions is necessary to more widely detect game-changing events in an eSports corpus. Lower inter-annotator agreement for the minority of event subtypes, however, suggests that further modifications should be made to event subtype

definitions.

There are multiple ways to continue improvement in event detection in eSports corpora. First, the *CS:GO* corpus could be lengthened to achieve a larger set of annotated events. This task could be accomplished efficiently with the addition of multiple transcribers of *CS:GO* broadcasts.

Second, event subtype definitions could be modified further. While the majority of game-changing events was detected in this corpus, others could be detected as well in a larger and more varied corpus. For example, bomb defusals are a round win condition for the counter-terrorist team in *CS:GO* and can have a major impact on gameplay. Creating an event subtype to detect less common but equally as important event types like bomb defusals can contribute to the detection of more game-changing events.

One advantage of conducting an artificial task such as this on an eSports corpus is that more clear-cut relations between events can be detected. While this paper specifically focused on event detection in the context of a *CS:GO* game, there is the possibility of using these event subtype definitions for annotation tasks of real-life events. This could prove especially successful in any corpora comprised of military texts where, what are considered game-changing events in a *CS:GO* game actually occur in real-life scenarios.

Ultimately, with extensive further development this corpus could be used as training data for an automated event detection system. A gold standard corpus could be produced not only for automation but also as the standard for the creation of further eSports corpora.

5 Acknowledgments

The author would like to thank Matthew Shanks for completing the annotation task, Auguste Massonant and Anders Blume for permission to use their speech as the basis of the *CS:GO* corpus, and Eileen Fitzpatrick for advice and guidance.

References

- American National Corpus Project. MASC - The Corpus Web. 9 March 2015.
- Donghun Lee and Linda J. Schoenstedt. 2011. Comparison of eSports and Traditional Sports Consumption

- Motives. *ICHPER-SD Journal of Research*. 6(2), 39-44.
- International Computer Science Institute. FrameNet Data Web. 17 April 2015.
- Jin-Dong Kim, Tomoko Ohta, Yuka Tateisi, and Jun'ichi Tsujii. 2003. GENIA corpus - a semantically annotated corpus for bio-textmining *Bioinformatics*. 19(1), 180-182.
- Linguistic Data Consortium. 2005. ACE (Automatic Content Extraction) English Annotation Guidelines for Events Version 5.4.3 2005.07.01. Web. 9 March 2015.