From the sidelines to the archived web: What are the most annoying football phrases in the UK?

Helena Byrne
Curator of Web Archives
@HBee2015
06/05/2019
Football Clichés

• Clichés have been a longstanding part of football

• In recent years linguistics of football has become an area of interest for academic study

• It is also the subject of some popular books

• Increase in media time to discuss football has led to an increase in the use of clichés in commentary

• This increase in use has also led to a dislike for this type of commentary and resulted in lots of most annoying phrases lists online through mainstream media and online forums
Data Collection

- Reviewed the lists on the first page of a Google search
- Conducted a survey from my social media accounts

<table>
<thead>
<tr>
<th>Men’s phrases reviewed</th>
<th>Women’s phrases reviewed</th>
<th>Total number</th>
</tr>
</thead>
<tbody>
<tr>
<td>343</td>
<td>123</td>
<td>466</td>
</tr>
</tbody>
</table>
### Most annoying phrases in football

<table>
<thead>
<tr>
<th>Men’s Phrases</th>
<th>Number of appearances</th>
<th>Women’s Phrases</th>
<th>Number of appearances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variants of comparing to players like Messi, Ronaldo etc.*</td>
<td>5</td>
<td>Boring</td>
<td>5</td>
</tr>
<tr>
<td>110 per cent</td>
<td>4</td>
<td>Lesbians</td>
<td>4</td>
</tr>
<tr>
<td>He’s hit it too well</td>
<td>4</td>
<td>Slow</td>
<td>4</td>
</tr>
<tr>
<td>It's a game of two halves</td>
<td>4</td>
<td>Ladies football</td>
<td>4</td>
</tr>
<tr>
<td>It could be a cricket score</td>
<td>4</td>
<td>Women’s football</td>
<td>4</td>
</tr>
</tbody>
</table>

* For the purpose of this review only a comparison to Ronaldo was used as he spent some of his career in the English Premier League and is more popular than Messi on the archived web: https://www.webarchive.org.uk/shine/graph?query=%22Cristiano+Ronaldo%22%2C+%22Lionel+Messi%22&year_start=1996&year_end=2013&action=update
Women’s experience in football

“virtually replicates their more general experience – that they have been treated as not only different to men but also inferior in many respects”

**Most annoying phrases in football**

<table>
<thead>
<tr>
<th>Men’s Phrases</th>
<th>Number of appearances</th>
<th>Women’s Phrases</th>
<th>Number of appearances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variants of comparing to players like Messi, Ronaldo etc.*</td>
<td>5</td>
<td>Boring</td>
<td>5</td>
</tr>
<tr>
<td>110 per cent</td>
<td>4</td>
<td>Lesbians</td>
<td>4</td>
</tr>
<tr>
<td>He’s hit it too well</td>
<td>4</td>
<td>Slow</td>
<td>4</td>
</tr>
<tr>
<td>It's a game of two halves</td>
<td>4</td>
<td>Ladies football</td>
<td>4</td>
</tr>
<tr>
<td>It could be a cricket score</td>
<td>4</td>
<td>Women’s football</td>
<td>4</td>
</tr>
</tbody>
</table>

* For the purpose of this review only a comparison to Ronaldo was used as he spent some of his career in the English Premier League and is more popular than Messi on the archived web: https://www.webarchive.org.uk/shine/graph?query=%22Cristiano+Ronaldo%22%2C+%22Lionel+Messi%22&year_start=1996&year_end=2013&action=update
Welcome
Welcome to SHINE, our prototype historical search engine.

About SHINE
This tool has been developed as part of the Big UK Data Arts and Humanities project funded by the AHRC. Read more about the project on our blog.

The data was acquired by JISC from the Internet Archive (IA) and includes all .uk websites in the IA web collection crawled between around 1996 until April 2013 (you can read more details about this dataset here). The collection comprises over 3.5 billion items (urls, images and other documents) and has been full-text indexed by the UK Web Archive. Every word of every website in the collection can be searched for and analysed.

What can I do?
Trend analysis
Use 'trends' to analyse the number of pages a word or phrase appears in the collection over a given period (within 1996-2013). Comparisons can be drawn by adding several words or phrases separated by a comma. E.g. cat, dog, goldfish
JISC UK Web Domain Dataset (1996-2013)

Number of Resources by Year

.uk resources archived by the Internet Archive 1996 to April 2013

Number of resources per year in the Shine data set https://britishlibrary.typepad.co.uk/webarchive/2016/02/updating-our-historical-search-service.html
“Football”, “Soccer”

Found 100 samples matching "Soccer" from 2008.

Matching Text | Link
--- | ---
charger 3.75 psp - fta | wholesalepages.co.uk
striking payments in hong kong. your | glendasgoriff@hotmail.co.uk
soccer 22.99 (£22.99) for 100+ | thegrieve.co.uk
pap - sandisk 1 gig mem cards 56.99 Comp... | thegrieve.co.uk
soccer is interdependent to me. Decide on your goals and talk them | goalsfootball.co.uk
Photo album > If you have any interesting photos then send them by n SOCCER SQUAD Trophy winners- pictures Ajax - Picture from the early days... | centralajax.co.uk
keygen. evolution | thegrieve.co.uk
soccer | thegrieve.co.uk
dog clip, safe n sec torrent, warcraft data maps for wind, new star | goalsfootball.co.uk
Home About Facilities Virtual Tour Corporate Branches News Main Net | goalsfootball.co.uk
¢ Goals | goalsfootball.co.uk
Soccer Resort UK 5's | goalsfootball.co.uk
Soccer Centres pl. 2005 | goalsfootball.co.uk
Privacy Policy | goalsfootball.co.uk
Corporate Site Map Website de... | goalsfootball.co.uk
Free DVDs Free tickets to upcoming music gigs Free Theatre Ticket | ukt-freestuff.co.uk
" PolyPhonic Ringtones » News » Popular » Alternative » Anthems » C | absolutelyringtones.co.uk
Soccer > Spanish » St. Nicholas » Swedish » Turkish » TV & Movie » Valen... | exportbettingadvice.co.uk
Betting School Betaginl Soccerstats Punters Lounge Sports Punter | omagh.gov.uk
Soccer Punter UK Betting Systems | omagh.gov.uk
Programmes & Events Directory Of Sport Sport & Leisure In Omagh... | omagh.gov.uk
- Soccer - Softball - Speleology - Squash - Strength_Sports - Table_... | gifts-occasions-4u.co.uk
soccer 2008 wii wii bowling for columbia cvd pre bowling for columbia (a... | oise.co.uk
Home | Create Account | Get Started Now! | Player Support | Help Spa | Futures | Soccer | Auto Racing | Golf | Boxing | General... | sports-betting-lines.co.uk
History Events Odds Scores Teams Players Informations Sports, Hocki... | sports-betting-lines.co.uk
| - NBA Hockey - NHL NFL Football | soccer | Football Betting Formula One | soccer | Auto Racing | Golf | Tennis

https://www.webarchive.org.uk/shine/graph?query=%22Football%22%2C+%22Soccer%22&year_start=1996&year_end=2013&action=update
"Soccer", "Football" NOT "Gaelic Football" NOT "Rugby Football" NOT "American Football" NOT "Aussie Rules Football"
A and B side by side

A = 15,843,947
B = 15,048,344
2010 Total = 226,639,215
"women's football", "ladies football", "girls football", "boys football", "men's football"
What are the most annoying football phrases on the UK archived web space?

"A game of two halves" AND "Football"
What are the most annoying football phrases on the UK archived web space?

"women's football", "Ladies Football", "Slow"

1996  2013

Update Graph  Reset All

https://www.webarchive.org.uk/shine/graph?query=%22women%27s+football%22+AND+%22Ladies+Football%22+AND+%22Slow%22+AND+%22Women%E2%80%99s+Football%22+AND+%22Ladies+Football%22+AND+%22Lesbian%22+AND+%22Women%E2%80%99s+Football%22+AND+%22Ladies+Football%22+AND+%22Boring%22+AND+%22Women%E2%80%99s+Football%22+AND+%22Ladies+Football%22&year_start=1996&year_end=2013&action=update

"women's football"
What are the most annoying football phrases on the UK archived web space?

“A game of two halves”

“Women’s football”
Main Issues

- Duplication
- A single host that spikes results
- Some phrases either too generic or too unique
- Football is a designated tab on many websites which can skew relevant results returned if the phrase is used outside of football.
However ….

- This dataset can still give us a general overview of how popular certain terms are over this period in time.

- As long as the search parameters are clearly outlined then the results to a search query should be seen as valid.
Useful Links ....

webarchive.org.uk

webarchive.org.uk/blog

webarchive.org.uk/videos

webarchive.org.uk/shine

data.webarchive.org.uk/opendata

web-archivist@bl.uk
@UKWebArchive

helena.byrne@bl.uk
@HBee2015