The Transcription and Representation of Spoken Political Discourse in the UK House of Commons

Cribb, M. and Rochford, S.

The Transcription and Representation of Spoken Political Discourse in the UK House of Commons

V. Michael Cribb[1] & Shivani Rochford[2]

[1] School of Humanities, Faculty of Arts and Humanities, Coventry University, United Kingdom

[2] Centre for Academic Writing, Coventry University, United Kingdom

Correspondence: Michael Cribb, School of Humanities, Faculty of Arts and Humanities, Coventry University, United Kingdom, Tel: 44-024-7765-8011. E-mail: aa6177@coventry.ac.uk

Abstract

This paper considers the representation of spoken political discourse through the transcription practices of the House of Commons. It considers how the Hansard method of transcription represents the oral debates during the weekly parliamentary sessions of Prime Minister's questions in three areas: lexical and grammatical fidelity, performance characteristics and interruptions from the audience. The paper also considers how accurately and faithfully speakers in the House quote from Hansard during these sessions as they pursue their arguments. The findings suggest that while Hansard does what it purports to do, modern transcription methods and digital representations necessitate additional tools to augment this system. We argue the case for a more representational, multi-tool approach to the transcription of discourse.

Keywords: political discourse, transcription, Hansard, representation, quotation, prime minister's questions

1. Introduction

Speech has always been a notoriously difficult medium to capture permanently. Written language by nature is there on the page for all to see and read; speech by comparison is transitory and lacks permanence. As soon as an utterance leaves the mouth, the message fades rapidly and listeners do not have the means of being able to rewind to listen again. Modern recording techniques sometimes enable a portion to be reviewed but it remains the case that the vast bulk of speech that reaches our ears fades very rapidly. We retain the "gist" of what is said but discard the verbatim record very soon after hearing it. It is important then that we employ robust procedures for transcribing and capturing speech in written form if we are to ensure a full representation of discourse is obtained.

The UK House of Commons is a chamber in which debates and discussions take place largely in an oral manner on a daily basis between elected Members of Parliament (MPs) of various parties. The spoken discourse generated via these debates is recorded daily in written form in the official transcripts of the House known as "Hansard". This is a "verbatim" report of what has been said in parliament but which removes from the transcription "repetitions and obvious mistakes" which are not recorded (Hansard Online, 2016). The purpose of our paper is to review this method of transcription to assess how faithfully it represents what was actually said and the discourse that took place. This is not meant in any way to be judgemental of Hansard, which has served the public well for many decades, as a record of parliamentary proceedings. Rather we look at ways in which this method of recording can be augmented given modern digital representations and highlight the nature of spoken discourse in a political setting. Two main questions are posed: whether Hansard under-represents the spoken discourse of the House of Commons and whether the use of Hansard to quote previous MPs leads to a fair representation of the original statement. The study looks at the discourse of debates during the weekly sessions of Prime Minister's questions which has become an important event in the British political calendar. Since commentators, journalists and broadcasters often rely on Hansard as a primary source for what has been said during these sessions, it is important that we understand how this written record has been constructed and its limitations.

2. Background

Writing developed circa 5000 years ago (Gross, 2012) and since then the medium has been used to transcribe

speech in various guises. With the invention of audio recording techniques, transcription has proliferated leading to the creation of a plethora of transcription records and databases for a wide range of speech events including telephone dialogues, dinner table conversations and political interviews to name just a few. The standard transcription method in the Applied Linguistics discipline today is Conversation Analysis developed by Sacks, Schegloff & Jefferson (1974) in which a full and faithful turn-by-turn transcription of a speech event is generated through a systematic and repetitive analysis of a recording of the original spoken discourse. This method attempts to capture everything that was verbalised including speech hesitation phenomena such as "ums" and "ahs" together with paralinguistic and non-verbal features where significant. This denaturalised transcription contrasts with more naturalised conventions, often employed outside the applied linguistic tradition, in which "the process of transcription is made less visible through literacization, the privileging of written over oral discourse features" (Bucholtz, 2000, p. 1461). Many researchers employ a transcription practice somewhere in between these two extremes however.

The transcription and subsequent analysis of spoken discourse are important procedures not only for the linguistic researcher but also for the professional working in the field. There have been many cases, for example, where poor transcription in criminal proceedings, or poor understanding of transcriptions methods, has led to the misrepresentation of statements by lawyers and juries (Coulthard, 2005). Transcription can help the professional in two ways. First, it can help them to understand what has been said. The deliberate act of repeatedly listening to a recording of speech, slowing it down and capturing all the elements of the verbal stream, can help the analyst to understand meanings and structures within the discourse. Second, transcription provides the professional with a permanent record in written form of what has been said which then enables them to move back and forth between recording and transcription as they utilise the resource. One crucial point to make here is that the original recording should not be separated from the transcript once made and then discarded (Hutchby & Wooffitt, 2008, p. 70). The purpose of transcription is not merely to replace one medium (spoken) with another (written). There has been an almost unlimited number of instances whereby transcripts of spoken discourse have been published in books, articles or online, and in so doing the link with the original recording is broken. Once the separation has taken place, there is little chance of restoring the link and the original recording is often lost to the reader forever. It is a sad fact that we live in a world where written transcripts of spoken discourse abound but are divorced permanently from the original recordings so that whoever happens to read the transcript is unable to recover the original. In this paper, we have attempted to provide live links to all samples of discourse presented to address this situation.

2.1 Hansard

Hansard is the official record of what is said in the British houses of parliament. The record has a long history but entered regular publication through the printer Thomas Curson Hansard in the 19th century (Ralphs, 2009). It has been produced regularly since for debates and discussions in both Houses of Parliament (Lords and Commons) with the printed version normally appearing the morning after a day's sitting. The official Hansard site states that Hansard is a "verbatim report of what is said in parliament" (Hansard Online, 2016). The site makes clear that Members' words are "edited to remove repetitions and obvious mistakes, albeit without taking away from the meaning" and is an indication that Hansard does not pretend to be what it is not—i.e., a full Conversation Analysis transcription of speech. The transcripts are made live by a number of reporters working in shifts and these transcripts are then handed on to editing staff who compile the day's report to be published the next morning (Mollin, 2007). The transcripts are thus "thrice removed from the original" (Mollin, 2007, p. 189).

Slembrouck (1992) suggested that the process of producing the Hansard report in this way acts to embody "institutional assumptions about what is more or less important in parliamentary discourse representation" (1992, p. 117). The "writtenness" of the record acts to filter out the "spokenness" of language and targets "explicitness" and "well-formedness" of expression (1992, pp. 104-106). He suggested that the process of recording parliamentary discourse "entail[s] a reduction of utterances to ideational claims and positions... leaving interpersonal and textual meanings [out]" (1992, p. 109). This "premium on ideational meaning" reflects a "representation culture" (1992, pp. 108-109) according to Slembrouck which enables institutionalised ideology and power to be manipulated through the use of language.

Mollin (2007) found that the Hansard report tended to omit "performance characteristics of spoken language" (2007, p. 193) such as false starts, repetitions and reformulations as well as "parliamentary speech" that is characteristic of the specific setting (such as the cry to "order!" from The Speaker). These modifications, together with substantial changes to lexical and grammatical features, ostensibly to make the text more formal, mean that Hansard, she suggests, needs to be "treated with caution if they are to be used for linguistic analysis" (2007, p. 209). According to Mollin, Hansard is a decontextualised, factual and formal report on the proceedings

in Parliament that is a "stronghold of conservative and formal British writing style" (2007, p. 208) but which is a "nightmare" (2007, p. 208) as a resource for linguistic analysis.

Members of the Houses of Parliament however have no such qualms about using Hansard as a resource. Antaki & Leudar (2001) have noted how it can be recruited by Members as an "unimpeachable source of parliamentary record" (2001, p. 467) to promote their own arguments. Thus a speaker can use Hansard to quote another Member's words said previously to disarm any objection to a position and as a "hostile, unchallengeable, proof of the rightness of the speaker's case" (2001, p. 486). Indeed the verb "to hansardise" has entered into the lexicon of the Houses of Parliament specifically to refer to the act of quoting a Member's words against themselves (2001, p. 477).

2.2 Prime Minister's Questions

Prime Minister's questions (PMQs) is a weekly event in the House of Commons in which the Prime Minister of the day attends the chamber to answer questions from Members on all sides. This can be a hostile affair, particularly when the Prime Minister (PM) receives questions from the Leader of the Opposition (LO). The LO is the given the opportunity to ask a series of challenging and probing questions of his or her own choosing to the PM, and the exchange is invariably accompanied by noise, heckling and interjections from backbench Members in support of their leader and party. There has been some criticism that the event in recent years has degenerated into a "rowdy, mud-slinging spectacle catered more towards shallow political point scoring than serious scrutiny of prime ministerial activity" (Bates, Kerr, Byrne, & Stanley, 2012, p. 2). Despite this criticism, the event has taken on even greater significance in the British political calendar and is regularly televised live on the main broadcasting channels along with subsequent commentary and analysis. Questions of how well a leader performed and "who won" are often asked and discussed, and the newspapers the following day invariably reflect the main talking points of PMQs in their headlines. Indeed, the quality and potential of a leader can be significantly enhanced or undermined through their spoken performance at PMQs. This has been illustrated recently as new leaders Jeremy Corbyn (Labour) as LO and subsequently Theresa May (Conservative) as PM have taken up prime roles at PMQs. Both leaders have had their performances at this event heavily scrutinised by the media and their character as leaders, and sometimes the fortunes of their party, is partly reliant on this scrutiny.

3. Methods

The present study aims to investigate the recording of discourse at PMQs through the Hansard report. Two main research questions are posed:

RQ1. Does Hansard under-represent the spoken discourse of the House of Commons during PMQs?

This question looks at how well the recording by Hansard represents the spoken discourse of the PMQ debates as they happened. Does the report accurately represent what has been said? What is not captured by the report?

RQ2. Do speakers at PMQs who rely on Hansard to quote other Members of the House accurately and faithfully represent what has been said previously?

Speakers at PMQs sometimes quote Hansard to highlight what other Members of the House have said often for the benefit of their own argument. If Hansard is under-representing the original speech, could quoting from it lead to a magnification of the misrepresentation? Do Members manipulate this process to promote their own argument?

3.1 Data

Data for RQ1 was taken from ten parliamentary sessions of Prime Minister's Questions (PMQs) over the period June 2012 to July 2014. PMQs was chosen as the focus rather than the whole of the House's business due to the importance of this event in the weekly political calendar in its ability to mark a leader's quality as a speaker. For each session, one question-answer exchange between the LO (Edward Miliband) and the PM (David Cameron) was selected together with one question-answer session between a backbench opposition MP and the PM. The LO-PM exchanges typically generate more adversarial discourse than backbench MP-PM exchanges although this is also dependent on the topic of the exchange. This selection process generated 20 question-answer exchanges in total. These are shown in table 1 with the topic of discussion.

Table 1. Question-answer exchanges at PMQs

Sample no.	Exchange with PM	Date of session	Topic
1	LO	13 Jun. 2012	Culture Secretary
2	Pat McFadden (Lab)	13 Jun. 2012	Leveson Inquiry and FSA
3	LO	24 Oct. 2012	Energy & Railways
4	Gavin Shuker (Lab)	24 Oct. 2012	NHS
5	LO	16 Jan. 2013	European Union
6	Grahame Morris (Lab)	16 Jan. 2013	European Union
7	LO	24 Apr. 2013	NHS
8	Nick Smith (Lab)	24 Apr. 2013	Care Homes
9	LO	10 Jul. 2013	Party Donations
10	Helen Goodman (Lab)	10 Jul. 2013	Banking Reforms
11	LO	27 Nov. 2013	State intervention in markets
12	Dianne Abbot (Lab)	27 Nov. 2013	London Underground
13	LO	15 Jan. 2014	Bank Bonuses
14	Tom Watson (Lab)	15 Jan. 2014	Amritsar Inquiry
15	LO	30 Apr. 2014	Sale of Royal Mail
16	George Howarth (Lab)	30 Apr. 2014	Debt and Poverty
17	LO	25 Jun. 2014	Leveson Inquiry
18	Ronnie Campbell (Lab)	25 Jun. 2014	Andy Coulson
19	LO	16 Jul. 2014	Economy and Employment
20	Huw Irranca-Davies (Lab)	16 Jul. 2014	Royal Mail

Note. PM=David Cameron; LO=Edward Miliband.

Data for RQ2 was taken from PMQs over the same period. However since Members don't quote from Hansard every session it was not possible to match this data with the dates in table 1. Instead six occurrences were selected from either the PM or LO as a representative sample (see table 2). Some of these quotes refer to speakers at PMQs but in other cases they refer to sessions outside of PMQs. In all cases however the sessions were business carried out in the Houses of Parliament and so Hansard records exist.

Table 2. Quotations from Hansard at PMQs

Sample no.	Speaker	Date of quote	Original speaker (and session)	Date of original	Topic
21	PM	9 Jul. 2014	PM at PMQs	2 Jul. 2014	NHS
22	PM	25 Jun. 2014	Chris Bryant in Westminster Hall debate	13 Nov. 2012	Andy Coulson
23	PM	30 Oct. 2013	LO at PMQs	23 Oct. 2013	Energy Markets
24	LO	7 May 2014	Vince Cable during Qs to Sec. of State	6 May 2014	Takeovers
25	PM	30 Oct. 2013	Lord Donoughue in House of Lords	28 Oct. 2013	Energy Prices
26	PM	17 Jul. 2013	LO at PMQs	18 Jan. 2012	Unemployment

3.2 Data Transcription

A full Conversational Analysis (CA) using the Jefferson notation (Jefferson, 2004) provides for a very detailed yet time-consuming method of transcribing spoken discourse. In this paper, we employ a modified, simpler form in which the major features of conversation such as pausing, hesitation and repetition are included as well as significant audience features as an illustration of what Hansard could aspire to. A transcription key is given in table 3. Orthographic transcription was maintained throughout except for very common transcription representations such as "gonna" for "going to". Each questions-answer sessions was transcribed from the audio archive of the session at parliamentlive.tv <http://www.parliamentlive.tv/Commons>. This led to a database of approximately 4,840 words representing just over 36 minutes of debate.

Table 3. Transcription key

Symbol or example	Notes
(.)	Micropause
(6.0)	Pause in speaker's words (seconds)
"er", "erm"	Hesitation phenomena restricted to two variants
"mister"	"Mr"-Lexical items spelled out in full
((laughter))	Feature of chamber in double braces. (See also table 13.)
xx	Word unclear

For the purposes of this paper, the Hansard report <https://hansard.parliament.uk/> and the CA transcript are positioned side by side in tabular form as shown in table 4. The Hansard text has been "expanded" where necessary to align it with the CA transcript as closely as possible. The CA transcription illustrates typical features which are not present in the Hansard report such as ((laughter)) from the chamber in line 05 and the eight second pause in the speaker's words at line 11 which was filled with ((general noise)) from the chamber. The transcription also highlights the difference between what was actually uttered by the speaker and what is reported in Hansard, a point which we will return to later. (A hyperlink link to the original audio-video recording is given below each sample.)

Table 4. Example of Hansard report and CA transcript (sample 1)

	Hansard report	CA transcription
01		((general noise))
02	Edward Miliband: I have to say that the right hon.	EM: I have to say (.) he's reached
03	Gentleman has reached a new state of delusion	a new state of delusion
04		I mean really and tru- hh
05	—really and truly.	((laughter))
06		really and truly
07	He just wants to talk about the past	you know (.) you know mister speaker
08		he just wants to talk about the past
09		he was-
10		((general noise))
11		(8.0)
12		he just wants to talk about the past
13		mister speaker

Source: http://parliamentlive.tv/event/index/e3e03421-9d26-4bee-bfb1-d39eabae05f0?in=12:10:15&out=12:10:36

4. Analysis

The analysis of the differences between the Hansard report and the CA transcription is reported here in three areas:

1). Lexical and grammatical changes to speech

2). Performance characteristics of the current speaker (e.g. hesitations, speech errors)

3). Interruptions from the chamber

4.1 Lexical and Grammatical

Table 5 show the percentage of changes to lexical and grammatical features within the Hansard report. The table compares the features that Mollin identified in her 2007 paper with the present study. These are all features which act to make the Hansard report more "written and formal in style" according to Mollin (2007, p. 208). The results here show that, for many features, the percent of cases of change by the Hansard editors has decreased. In the features where the percentage has increased (feature 7, 10, 15), the low occurrence of that feature in our data means the increase is not statistically significant. The general decrease in the changes could imply that the Hansard editors have moved towards a less formal, less conservative representation of spoken language. (We should also note that Mollin considered different types of parliamentary sessions whereas the present study only considers PMQs. It could be that the scribes are less willing to alter spoken features at PMQs due to the significance of the event in the public's eye.)

Table 5. Comparison with Mollin 2007

No.	Feature	Mollin 2007		Present study	
		Percent of cases	Type of change	Percent of cases	Type of change (occurrences if low)
1	generic *you*	100%	omitted or changed to impersonal and third-person constructions	70%	omitted or changed to impersonal constructions
2	*government* with singular concord	100%	omitted or changed to plural concord	100%	changed to plural concord (1 occurrence)
3	*help someone* with bare infinitive	100%	omitted or changed to *to*-infinitive	n/a	(no occurrences)
4	*Speaker*	100%	omitted	92%	as Mollin
5	contractions	98.7%	omitted or changed to non-contracted forms	96%	as Mollin
6	*give way*	97.7%	omitted	n/a	(no occurrences)
7	*going-to*-future	93.8%	omitted or changed, esp. to the *will*-future	100%	changed to *will* (2 occurrences)
8	*absolutely*, *actually*, *clearly*, *really*, *very*	75.0% (ave.)	omitted or changed	10%	omitted
9	determiner *this/these*	72.8%	omitted or changed, esp. into *that/those*	22%	as Mollin
10	*MAKE sure*	70.8%	omitted or changed, esp. into *ensure*	100%	changed to *ensure* (2 occurrences)
11	*LOOK at*	64.3%	omitted or changed, esp. into *consider* and *examine*	0%	(1 occurrence)
12	*order*	53.9%	omitted	44%	as Mollin
13	first person singular pronouns	45.9%	omitted	16%	as Mollin
14	*a number of*	40%	omitted or changed, esp. into *several*	0%	(1 occurrence)
15	*have to*	39.0%	omitted or changed, esp. into *must*	67%	omitted (3 occurrences)

One of the biggest changes observed in the present study compared with Mollin is in feature 1 (generic "you"). Mollin noted that this was almost invariably changed to either an "impersonal structure or construction with a third person pronoun" (Mollin, 2007, p. 204). The editors these days are more inclined to retain the generic "you" as in table 6 in line 01.

Table 6. Retention of generic "you" (sample 4)

	Hansard report	CA transcription
01	The Prime Minister: What is true is you	PM: what what is true is you
02	can always guarantee that Labour Members	can always guarantee that labour members
03	of Parliament will get up in Parliament and	of Parliament will get up in Parliament and
04	scaremonger about our NHS.	scaremonger about our N-H-S

Source: http://parliamentlive.tv/Event/Index/974eccc9-36ca-4fc6-a15c-e61772ccc7ad?in=12:16:52&out=12:17:00

Hansard is also more likely to retain feature 8 (amplifiers and stance adverbials) nowadays. Only 10% of cases were omitted in our data compared with 72.84% at the time of Mollin's study. One rare case of omission of the adverb "clearly" is given in Table 7 in line 02.

Table 7. Omission of adverbial (sample 16)

	Hansard report	CA transcription
01	The Prime Minister: What I would say to the	PM: wh- what I would say to the
02	right hon. Gentleman is that the best	honourable gentleman is (.) clearly the best
03	route out of poverty is work	route out of poverty is work

Source: http://parliamentlive.tv/Event/Index/c7e061a5-7c42-4f59-b5c4-4a91488efd25?in=12:38:51&out=12:40:00

Another significant change in the report is with feature 9—determiners *this/these*—which Mollin found were frequently changed to *that/those* or omitted, ostensibly she suggests to "remove the speech from the here and now" (Mollin, 2007, p. 200). In our data, the percentage occurrence of this change has dropped to 22%. An example of a change is given in table 8 in line 04.

Table 8. Change of determiner (sample 8)

	Hansard report	CA transcription
01	The Prime Minister: ... One of the most	PM: ... I think one of the most
02	important things we can do is ensure that the	important things we can do is (.) is make sure the
03	Care Quality Commission is up to the task of	care quality commission is up to the task (.) of
04	investigating those homes properly and has	investigating these homes properly (.) and has
05	robust structures in place	really robust structures in place

Source: http://parliamentlive.tv/Event/Index/615e9b50-cba6-4006-9ebe-596341d17c79?in=12:21:04&out=12:21:16

The results above suggest a move towards a less formal more faithful-to-the-original style of reporting of parliamentary discourse but Hansard still makes significant changes and omissions to lexical and grammatical features in other areas. In Table 9, "the financial services act" in line 03 is changed to just "the Act" and "he talks about" in line 04 is changed to "the right hon. Gentleman mentions". The full name of the Act is "the Financial Services and Markets Act", which had been mentioned by the previous speaker, and it could be the editors in making this change do not want the record the Prime Minister as having used a reduced form of the name.

Table 9. Changes and omissions of lexical and grammatical features (sample 2)

	Hansard report	CA transcription
01	The Prime Minister: Clearly there are very	PM: Clearly there are very
02	strict rules, including the stock exchange	strict rules, including the stock exchange
03	code and the Act that the right hon.	code (.) and the financial services act that
04	Gentleman mentions	he talks about

Source: http://parliamentlive.tv/Event/Index/e3e03421-9d26-4bee-bfb1-d39eabae05f0?in=12:15:16&out=12:15:23

In the following example (table 10), multiple changes are made to Ed Miliband's original words. While the gist of what is said is retained the surface form has been significantly altered.

Table 10. Changes and omissions of lexical and grammatical features (sample 15)

	Hansard report	CA transcription
01	Edward Miliband: We have had no answer	EM: n- no answer to the question mister
02	to the question, Mr Speaker. The Royal Mail	speaker (.) the royal mail share price is
03	share price is currently 50% above the level	currently fifty percent above the level it
04	at which it was sold. Only the Prime	was sold at (.) only he would want
05	Minister would want to be congratulated on	congratulation (.) for losing the taxpayer
06	losing the taxpayer £1 billion.	one billion pounds

Source: http://parliamentlive.tv/Event/Index/c7e061a5-7c42-4f59-b5c4-4a91488efd25?in=12:11:23&out=12:11:35

4.2 Performance Characteristics

All speech contains hesitations and disfluencies which are a natural part of the flow of words from the mouth. These phenomena consists of speech errors, repetitions and false starts, as well as filled and unfilled pauses (Lickley, 2015). Only under the most carefully controlled speaking conditions can a speaker avoid these features and in some cases they may be used strategically for effect. Hansard does little to capture these characteristics of speech, yet a speaker's performance in the House of Commons, especially during Prime Minister's questions, can be crucial to how they are perceived both from their peers and by the public at large. Careers can be made or lost by how well a speaker performs at PMQs particular leaders of the main parties such as the Prime Minister and the Leader of the Opposition.

The present study found that 11.7% of the lexical words spoken in the exchanges were omitted from the final Hansard publication. This excludes fillers such as "er" and "erm" and comments from the chamber, but includes

repetitions, false starts and all other utterances by the speaker. This figure compares with 18% for Mollin (2007, p. 192). This could indicate that the editors are including more of the original utterances in Hansard compared with 10 years ago which supports the suggestion in the previous section that the report has become less formal in its style and more faithful to the original spoken discourse. This is further supported when we consider that Mollin (2007) analysed a variety of sessions in the House of Commons most of which were more "sedate" affairs than the "rowdy" PMQs studied here. The nature of PMQs means that the discourse is likely to contain performance features which will be removed eventually from Hansard.

The adversarial and rowdy nature of the LO-PM exchanges can also be highlighted over the more sedate backbench MP-PM question-answer exchanges at PMQs where opposition Members only have one question often prepared in advance. The results reveal that 13.5% of verbal utterances were removed from PM-LO exchanges compared with 8.4% from exchanges between the PM and backbench opposition Members. This difference is probably due to the higher degree of heckling and noise during PM-LO exchanges combined with more verbal sparring which leads to greater hesitation and floor holding rhetoric by the person at the dispatch box.

As an example of this during LO-PM exchanges, consider table 11. The chamber is particularly boisterous and noisy during Miliband's (LO) turn at the dispatch box and the strategy for Miliband is to hold the floor through repetitions and false starts. This is preferred option for most speakers when being heckled in the chamber rather than lengthy pauses (although note pauses in lines 11 and 13) which could lead to a loss of face and possibly an intervention by The Speaker. In the sample presented, nearly 50% of Miliband's spoken words do not find their way into the Hansard report.

Table 11. Holding the floor (sample 1)

	Hansard report	CA transcription
01		((general noise))
02	Edward Miliband: I have to say that the right hon.	EM: I have to say (.) he's reached
03	Gentleman has reached a new state of delusion	a new state of delusion
04		I mean really and tru- hh
05	—really and truly.	((laughter))
06		really and truly
07	He just wants to talk about the past	you know (.) you know mister speaker
08		he just wants to talk about the past
09		he was-
10		((general noise))
11		(3.0)
12		he-
13		(3.0)
14		he just wants to talk about the past
15	INSERT BLANK LINE HERE	mister speaker
16	—he was the future once.	he was the future once
17		((cheers and general noise))
18		now (.) and isn't the truth
19		and isn't the truth
20		and isn't the truth
21		((noise continues))
22	The Deputy Prime Minister	the deputy prime minister
23		the deputy prime minister
24		the deputy prime minister
25	says	says says the bid- sh-
26	that the decision should go to the independent adviser	says says the decision should go to
27		the independent adviser

Source: http://parliamentlive.tv/Event/Index/e3e03421-9d26-4bee-bfb1-d39eabae05f0?in=12:10:14&out=12:10:58

Similarly, in table 12 the Prime Minister rises amid general noise from the chamber to answer a question from Ed Miliband during a LO-PM exchange. Unusually for Hansard, the report indicates what was said by the chamber, "Hon. Members" in this case, in lines 03-04 as they demand the PM answers the LO's question. The PM holds the floor through hesitation devices and responds to the chamber in line 05 and 06.

Table 12. Holding the floor (sample 9)

	Hansard report	CA transcription
01		((general noise))
02	The Prime Minister: First of all, let	PM: first first first of all
03	me deal with—[Hon. Members:	let me deal with
04	"Answer!"]	let me deal
05	I will answer.	I will answer
06		I will answer
07	Let me deal with 6p a week—	let me deal with six p a week
08	[Interruption.]	((interruption from The Speaker))

Source: http://parliamentlive.tv/Event/Index/82d70f45-a0c2-4d83-b96a-f73a623cda42?in=12:06:57&out=12:07:09

While we are not arguing that all hesitation phenomena are included in Hansard, we do believe that they are important features of spoken performance and as such cannot simply be edited permanently from the record.

4.3 Interruptions

The rules of the House do not give the audience "ratified speaking rights" (Ikeda, 2009, p. 52) during PMQs but the event does not take place within a silent chamber. Indeed, one of the defining features of PMQs, and one of the reasons why it has such a reputation as a "bear pit" for politicians, is the invariable interaction between speaker and audience (i.e., the Members in the chamber). Because of the close physical proximity of the speaker at the dispatch box to Members of the chamber, and due partly to the adversarial arrangement which seats rival politicians directly facing each other, there is a great deal of cheering, barracking and heckling which contributes to the discourse and often dictates how proceedings unfold. Some commentators have likened the event to a rowdy football match in which each side has behind them a group of supporters who cheer and "sing" their side's praises while maintaining a "barrage of noise... designed to unsettle the other side" (Hoggart, 2011). The "power" of PMQs over the political status of politicians and parties is often established not through what was said but the reaction of the chamber to what was said as this is likely to dictate the headlines in the media the following day.

The most Hansard will usually do to represent audience-speaker interaction is insert the word "Interruption" into the transcript (e.g., line 08 in table 12) to indicate some sort of disruption to the flow of proceedings without specifying the nature of the disruption. However this one-size-fits-all word does not capture well the subtle complexities of the interaction between speaker and audience. Table 13 indicates a brief taxonomy of the type of interruptions that can occur during PMQs that have been observed in the data under consideration.

Table 13. Interruptions from chamber

Category	Description
((general noise))	A general increase in the level of noise in the chamber caused by many Members on all sides either talking or shouting.
((shouts of agreement))	General shouts and cheers of "yeh" or "hear hear" by many Members of one party to signal agreement with the current speaker.
((shouts of dismay))	General shouts of "no" or "ahh" by many Members of the opposite party to signal dismay at the current speaker's words or avoidance of topic.
((wooo!))	A collection of noises by several Members (often of one party) to signal surprise or shock at what has been said; often a "wooo" of surprise or something similar.
((yes-yes-yes chorus))	A series of three or more shouts of "yes" or other word by one party in rhythm and timed with the current speaker who is listing points; acts to emphasise the points and provide support.
((laughter))	General laughter by several Members at what has been said; may be accompanied by cheers; laughter may be genuine or mocking.
((revealing "ah"))	A collective expression of mock "realisation" by several Members of one party in response to a fact that has just been revealed.
((misc))	"hear hear" – by one or two Members laughter or noise – by one person other noises, grunts, groans that are picked up by the microphones.

Because of the power of the audience during PMQs, it is important that speakers, especially the Prime Minister and the Leader of the Opposition, rally their supporters around their line of attack and design their monologue in part for the audience (Schröter, 2014). Most of the support from Members consists of shouts of agreement for their party leader or shouts of dismay aimed at the opposition when they feel the speaker is equivocating or avoiding the topic altogether. However a number of other interesting interactions are evident in our data which reveal the intricate nature of the political discourse during PMQs.

One such interaction which was often employed to good effect by the former Prime Minister David Cameron and the Conservative backbench was a "yes-yes-yes" chorus (table 14). This is a collective series of "yes" by backbench Members timed to the current speaker's points to add emphasis to each one. The Hansard transcript simply records the Prime Minister saying "yes" three times but this belies the force with which it has been said (with stress and high falling intonation which is included in the CA transcript) and the timed and rhythmic echoing of this by his backbenchers. The Conservative Prime Minister here is clearly trying to paint the Labour party as being controlled by the Unions and by co-opting his supporters in this manner, whether staged or spontaneous, he makes best use of prime-time television coverage to convey his point.

Table 14. Yes-yes-yes chorus (sample 9)

	Hansard report	CA transcription
01		((general noise))
02	The Prime Minister: .. Will the unions	PM: ... will the unions still have
03	still have the biggest vote at the	the biggest vote at the conference
04	conference? Yes.	↘ yes
05		((chamber: "↘ yes"))
06	Will they still be able to determine the party's policy?	will they still be able to determine the party's
07	Yes.	policy
08		↘ yes
09	INSERT BLANK LINE HERE	((chamber: "↘ yes"))
10	Will they still have the decisive vote in choosing the	will they still have the decisive vote
11	Labour leader?	in voting for the Labour leader
12	Yes.	↘ yes
13		((chamber: "↘ yes"))
14	Those are the facts	that is the facts

Note. Underlining indicates stress.
Source: http://parliamentlive.tv/Event/Index/82d70f45-a0c2-4d83-b96a-f73a623cda42?in=12:11:30&out=12:11:46

Another audience feature is the "revealing ah" (table 15 line 09). This is a collective expression ("ah") by a group of backbencher Members in response to some fact which has just been revealed by the speaker about the opposition. This expression of surprise is entirely false since Members have either been warned in advance by party whips to expect the revelation or already know about the fact from the media. The aim of the expression is to mock the opposition (or opposition leader) and to add emphasis to the current speaker's "revealing" fact.

Table 15. Revealing "ah" (sample 3)

	Hansard report	CA transcription
01		((general noise))
02	The Prime Minister: ... When he	PM: ... when he was Energy Secretary
03	became Energy Secretary,	when he became Energy Secretary
04	the companies were making	the companies were making
05	a £25 loss per bill;	twenty five pounds (.) loss (.) per bill
06	when he left government,	when he left government
07	they were making	they were making
08	£55 profit per bill.	fifty five pounds profit per bill
09		((chamber: revealing ah))
10	He did not stand up to the vested	he didn't stand up to the vested interests
11	interests; he stuffed their pockets	he stuffed their pockets with cash
12	with cash.	

Source: http://parliamentlive.tv/Event/Index/974eccc9-36ca-4fc6-a15c-e61772ccc7ad?in=12:07:19&out=12:07:40

The revealing "ah" can be said by one Member but it is normally "sung" in unison by a number of Members. It can vary in length but is normally 1 to 2 seconds. The intonation pattern is invariably rise-fall which is a complex intonation pattern in English signalling something surprising but final. In other words, "the information is revealing but not unexpected". When uttering the "ah" the MP may also include a rise and fall of the head (gesture) to reinforce the rise and fall in intonation.

4.4 Quoting from Hansard

As Antaki & Leudar (2001) have suggested, quoting an opponent's words from Hansard provides the speaker with an "unimpeachable" source with which to support their position. One concern however is that the process of quoting parliamentary discourse goes through several layers of representation, interpretation and re-interpretation. There is the initial interpretation of the original Member's word by the Hansard reporter and its subsequent representation in the Hansard publication. This is then interpreted by the speaker and quoted in the House (to what extent the original words are listened to is not known). The Hansard editors then determine that this is a quote and verify its accuracy in Hansard; they do not automatically assume that the speaker's quote is accurate (Antaki & Leudar, 2001, p. 471). It then appears in Hansard in quotation marks, set off from the speaker's words, and with a link to the original report.

With such a convoluted system for speakers quoting other Members, one might suppose it is prone to abuse and misquotation. However, from the limited number of samples analysed here this does not appear to be the case. There were no examples of egregious misquoting or misrepresentation in our data and the system appears to be remarkably robust. Table 16 lists the quotes analysed and the speaker making the quotation (see Table 2 also). The third column indicates whether a misquotation of the original words was observed. In all quotes this was not the case.

Table 16. List of quotations analysed

Sample no.	Speaker	Misquotation?	Rhetorical device used (if any)
21	PM	no	none
22	PM	no	none
23	PM	no	Selective quoting.
24	LO	no	Selective quoting
25	PM	no	Quotation over long distance
26	PM	no	PM assigns prediction on unemployment to LO rather than OBR

One reason for the findings may be that the Hansard method for recording discourse, which relies on interpretation by reporters, could actually work in favour of fairness in this instance. The Hansard report becomes the "official" record of what was said in parliament and thus lessens the chance for Members to interpret the words of the opposition in the way they might hope to (unlike utterances made outside of Parliament). The power of the written word in effect is the deterrent for any misquotation. This does not mean of course that Members cannot use quotes selectively as rhetorical devices to bolster their own position, as Antaki & Leudar (2001) noted. Quoting normally involves selecting certain words and omitting others which are the prerogative of the person constructing the quote. As an example of how this works, consider sample 26 where the PM is quoting the LO. The quotation by the PM is given in table 17 and the original quote from the LO, a year and a half earlier, is given in table 18. The focus of consideration is the process by which the discourse in the right hand column of table 18 becomes represented in the left hand column of table 17.

Table 17. Quotation from Hansard by PM (sample 26)

	Hansard report	CA transcription
01		((general noise))
02	The Prime Minister: last year he	PM: last year he said this
03	said that "next year, unemployment	next year unemployment will get worse
04	will get worse, not better, under his	not better under his policies
05	policies.	((laughter))
06	Nothing that he can say can deny	nothing he say can deny that
07	that"—[Official Report, 18 January	
08	2012; Vol. 538, c. 739.]?	
09	Is it not time he withdrew that and	isn't it time to withdraw it
10	admitted he was wrong?	and admit you were wrong
11		((shouts of agreement))

Source: http://www.parliamentlive.tv/Event/Index/d8d5b1f5-c9e3-4492-9280-ebee99f934f2?in=12:09:54&out=12:10:07

The quotation is a fair representation of what was said in the sense that the speaker (PM) has quoted accurately from Hansard. The Hansard reporter has recognised this as a quote and placed it in quotation marks together with a citation of the Official Report (table 17 line 03-08). However from a rhetorical point of view, the PM has been astute to assign the claim that "unemployment will get worse" to the LO when in fact the LO was using figures from the OBR (Office for Budget Responsibility, line 03 table 18). The LO re-emphasises this in line 10 "according to the independent OBR". The PM reinterprets the ambiguity in "Nothing that he can say can deny that" (line 08 table 18) to imply that the LO's prediction cannot be denied whereas the original utterance referred to the OBR's prediction.

Table 18. Original words by LO

	Hansard report	CA transcription
01	Edward Miliband: The Prime	EM: mister speaker he doesn't seem to
02	Minister does not seem to	understand
03	understand. The reason why the OBR	the reason the OBR figures matter
04	figures matter is that they show that	is that they show over the next year
05	over the next year, unemployment	unemployment will get worse not better
06	will get worse, not better, under his	on his policies
07	policies.	an- and he can-
08	Nothing that he can say can deny	and nothing he can say (.) can deny that
09	that. That long list of policies,	that long list of policies
10	according to the independent OBR,	according to the independent OBR
11	will make no difference.	will make no difference

Source: http://parliamentlive.tv/Event/Index/c2995259-618b-4907-80d8-30689b1b5fa0?in=12:04:36&out=12:04:57

This example of how spoken words come to be represented in Hansard over time between the original statement and quotation is quite illuminating. It demonstrates on the one hand how Hansard, while under-representing a Member's words, acts to fix them in time and space thus preventing other Members in the future from misquoting them. (Woe betide a Parliamentarian who misquotes Hansard!) But on the other hand it provides Members with an "unimpeachable source" (Antaki & Leudar, 2001) from which to make mischief. (Informal sources suggest that Members of Parliament are able to request changes to Hansard if they feel their words have not been represented correctly which might suggest a good deal of time and resources are invested in scrutinising Hansard to make sure it is "bullet-proof" for the future.)

5. Discussion

Bucholtz (2000, p. 1441) following on from Green, Franquiz & Dixon (1997) distinguishes between transcription as an "interpretive process and ... as a representational process". The Hansard report, as a largely interpretive process, does what it "says on the tin" in that it records the words of MPs while removing "repetitions and obvious mistakes" (Hansard Online, 2016). Our analysis here is in no way meant to be a criticism of this role it plays in this respect, which has served the country well for many years. Like Mollin (2007, p. 208), we do not wish to "judge the work of the Hansard reporters and editors". However we would question whether one "tin" by itself is sufficient robust for a modern parliamentary democracy in which political discourse is valued so highly and quoted widely in the media. Does the process of "literacization" (Bucholtz, 2000, p. 1461) run the risk of ignoring the spoken form and the context in which it was uttered? Is there a case for additional tools which would augment Hansard and promote a more robust interpretation of political discourse? We believe there is and would suggest two additions to the procedure for the recording of spoken discourse in the House.

The first suggestion is to have a parallel Conversation Analysis (CA) transcript alongside the Hansard transcripts with hyperlinks between the two (for the online version). Any politician, journalist or researcher would then be able to read the original Hansard transcript and, if desired, jump immediately via a hyperlink to the more detailed CA transcript, although we recognise the problem of making the text less transparent for readers unaccustomed to CA features in written text (Bucholtz, 2000, p. 1461). The link between the two transcripts would need to be maintained at the level of speaker and possibly even sentence. In other words, a user of Hansard who was reading what a particular MP has said would need to be able to click and jump to the CA transcript of the particular MP. Any delay in this jump or mispositioning (say if a user is redirected to the start of the parliamentary session in the CA transcript rather than to the particular MP) would render the system unsuitable and would defeat the object of the system which is to allow the user to browse seamlessly back and forth

between transcripts. The advantages of maintaining two versions of transcripts, a "naturalized" and "denaturalized" version as Oliver, Serovich, & Mason (2005, p. 13) call them, include facilitation of "member-checking" and different "types of analyses".

The production of a parallel CA transcript would obviously induce a cost and could be limited initially to certain sessions of the House such as Prime Minister's questions where the discourse is under greater scrutiny. We would suggest that the level of detail provided in the CA transcripts here represents a good balance between time to produce and faithfulness to the original discourse. This would enable original lexical and grammatical features to be retained, performance characteristics to be included and significant interruptions and interjections from the chamber to be captured. This would therefore enhance the validity of Hansard as a system of reporting of Parliamentary discourse and leave it less vulnerable to the charge that it under-represents spoken discourse, giving it greater "trustworthiness and accuracy" (Markle, West, & Rich, 2011).

A second suggestion to improve the system of representation of political discourse is to have hyperlinks from Hansard to the original recordings of sessions, Any user reading Hansard should be able to click and jump to an audio (or video) recording of the particular MP. Again, the link between the two resources needs to be maintained at the level of speaker or even, more importantly in this case, utterance. Currently it is possible for a user to read Hansard on one website while listening to or watching a recording of the parliamentary session on another. However, the time required to set up this dual mode of analysis and the complexity of aligning the two records means that it is far from perfect and only the most persistent analyst can be successful (as we have discovered to our cost!). Anyone not familiar with this process will soon abandon it. For a dual system such as this to be useful, the user needs to be able to move back and forth seamlessly and effortlessly between transcript and recording at the level of utterance. As suggested earlier, a transcript of spoken discourse that is divorced from the original recording is no transcript at all and it is great shame that many disciplines, including the field of applied linguistics, pay scant attention to this detail. Modern digital representations and HTML hypertextual features means that the internet and online resources are in a pre-eminent position to solve this challenge. Markle, West & Rich (2011) recognise this when they suggest we are on the precipice of change in transcription as emergent technologies afford us the ability to embed audio and video data into our representations, thus avoiding "the loss of meaning and unavoidable interpretation bias inherent in transcription".

6. Conclusion

This paper has suggested that the Hansard report, while providing a robust reporting procedure for what has been said in the Houses of Parliament and serving the country so well over the years, does under-represent spoken discourse, which as linguistics we find difficult to accept. This under-representation is particularly noticeable in the lexical and grammatical changes to speech Hansard makes, as well as detail on speech performance characteristics and interruptions from the chamber which it largely omits. While we accept that Hansard does what it purports to do and what it was designed for, we suggest that in a modern parliamentary democracy where political discourse is valued so highly, and widely disseminated via the media, the system for recording MP's words needs to be augmented with additional features to enable a fuller scrutiny of what is said, how it is said and the reaction to what is said. We suggest two such tools could be a detailed parallel Conversation Analysis (CA) transcript in the applied linguistic tradition and an utterance-level hyperlink facility between the Hansard report and the audio recording of the discourse. The CA transcript should be closely aligned with the Hansard text and should enable researcher and professionals to move between the two resources seamlessly and effortlessly. Similarly the hyperlinks between Hansard report and the original audio recording should be closely aligned so that a user at a particular point in the report can link through to the audio recording effortlessly and accurately. The current system of accessing Hansard on one website and the recording on another does not facilitate this.

Our call here for these additional tools for representing spoken discourse is a call really to a much wider community, not just the editors of the Hansard report. With the spread of the internet and digital representations, society is relying more and more on written representations of spoken discourse, not only in the political field but in other areas such as the legal profession and applied linguistics. It is vital that we recognise that written representations (i.e., transcripts) are just that: "representations", and not replacements for the original discourse. The onus for the provision and maintenance of the link between the transcript and the original recording is with the creator of the database, the transcriber, not the reader.

References

Antaki, C., & Leudar, I. (2001). Recruiting the Record: Using Opponents' Exact Words in Parliamentary Argumentation. *Text, 21*(4), 467-488. https://doi.org/10.1515/text.2001.008

Bates, S. R., Kerr, P., Byrne, C., & Stanley, L. (2012). Questions to the Prime Minister: A Comparative Study of PMQs from Thatcher to Cameron. *Parliamentary Affairs, 67*(2), 253-280. https://doi.org/10.1093/pa/gss044

Bucholtz, M. (2000). The Politics of Transcription. *Journal of Pragmatics, 32*(10), 1439-1465. https://doi.org/10.1016/S0378-2166(99)00094-6

Coulthard, M. (2005). Some Forensic Applications of Descriptive Linguistics. *Veredas—Revista de Estudios Linguísticos, Juiz de Fora, 9*(1) 2, 9-28. Retrieved from http://www.ufjf.br/revistaveredas/files/2009/12/artigo016.pdf

Green, J., Franquiz, M., & Dixon, C. (1997). The Myth of the Objective Transcript: Transcribing as a situated act. *TESOL Quarterly, 31*(1), 172-176. https://doi.org/10.2307/3587984

Gross, M. (2012). The Evolution of Writing. *Current Biology, 22*(23), 981-984. https://doi.org/10.1016/j.cub.2012.11.032

Hansard Online. (2016). *About Hansard Online*. Retrieved from https://hansard.parliament.uk/about

Hoggart, S. (2011). Prime Minister's Questions—or an Unpleasant Football Match? *The Guardian*. Retrieved from https://www.theguardian.com/politics/2011/dec/14/prime-ministers-questions-football-match

Hutchby, I., & Wooffitt, R. (2008). *Conversation Analysis*. Cambridge: Polity Press.

Ikeda, K. (2009) Audience Participation through Interjection: Japanese Municipal Council Sessions. *Journal of Language & Politics, 8*(1), 52-71. https://doi.org/10.1075/jlp.8.1.04ike

Jefferson, G. (2004). Glossary of Transcript Symbols with an Introduction. In G. H. Lerner (Ed.), *Conversation Analysis: Studies from the First Generation* (pp. 13-31). Amsterdam: John Benjamins. https://doi.org/10.1075/pbns.125.02jef

Lickley, R. J. (2015). Fluency and Disfluency. In M. A. Redford (Ed.), *The Handbook of Speech Production* (pp. 445-469). London: Wiley Blackwell. https://doi.org/10.1002/9781118584156.ch20

Markle, D. T., West, R. E., & Rich, P. J. (2011). Beyond Transcription: Technology, Change, and Refinement of Method. *Forum: Qualitative Social Research, 12*(3). Retrieved from http://www.qualitative-research.net/index.php/fqs/article/view/1564

Mollin, S. (2007). The Hansard Hazard: Gauging the Accuracy of British Parliamentary Transcripts. *Corpora, 2*(2), 187-210. https://doi.org/10.3366/cor.2007.2.2.187

Oliver, D. G., Serovich, J. M., & Mason, T. L. (2005). Constraints and Opportunities with Interview Transcription: Towards Reflection in Qualitative Research. *Soc Forces, 84*(2), 1273-1289. https://doi.org/10.1353/sof.2006.0023

Ralphs, K. (2009). Recording Parliamentary Debates: A Brief History with Reference to England and New Zealand. *Australasian Parliamentary Review, 24*(2), 151-163.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974) A Simplest Systematics for the Organization of Turn-taking for Conversation. *Language, 50*, 696-735. https://doi.org/10.1353/lan.1974.0010

Schröter, M. (2014) Addressee Orientation in Political Speeches: Tracing the Dialogical 'Other' in Argumentative Monologue. *Journal of Language & Politics, 13*(2), 289-312. https://doi.org/10.1075/jlp.13.2.05sch

Slembrouck, S. (1992) The Parliamentary Hansard "Verbatim" Report: The Written Construction of Spoken Discourse. *Language and Literature, 1*(2), 101-109. https://doi.org/10.1177/096394709200100202