

Exploring a Learning Ecology: Teenagers' Literacy Practices in a Teen Second Lifeⁱ Project - Shome Park

Paper presented as part of a symposium: Researching the Literacy Practices of Children and Young People in Virtual Worlds, American Educational Research Association Conference, New Orleans, April 8-12, 2011.

Julia Gillen

Literacy Research Centre, Lancaster University, UK

This paper examines the connections made by participants in the Shome Park programme, an informal education project set in a virtual world, and their experiences of schooling. Learning is regarded as transformations in the patterns of participation in joint activity (Rogoff, 1997). Students and staff, while diverse in degree and duration of participation, all experienced the project as new in many respects. I took as my starting point for investigation the assumption that participants, all with ongoing involvement in other forms of education, made sense of an innovative learning environment in part through drawing on their experience of practices in other domains.

Exploration implies new territory; in this paper I endeavour to offer an innovative approach to the empirical investigation of a substantive question.

Methodologically, I introduce some techniques from corpus linguistics, as appropriate to the study of a huge volume of digitized texts. Substantively, I reveal a number of ways in which participants made links with their school-based identities, contrasted or compared their experiences across domains and demonstrated through their interactive written communications their 'learning as you go' participatory online culture (Lankshear and Knobel, 2006).

Methodologically, the paper demonstrates some possibilities of corpus and discourse analytic approaches to digital projects that generate vast data records.

This study and the experiences of the project more generally as reported elsewhere may contribute to our reaching understandings of how learning takes place across settings, and of the possible synergies and obstacles involved. This may be useful to educators interested in finding ways to supplement or extend school-based opportunities (Barron, 2006; Guo, Amasha & Tan, 2011).

The Schome Park programme (SPP) that forms the focus of this study was a thirteen-month long engagement by a virtual community working with teenagers located in the UK and the USA. Most participated outside their school environments but SPP included after school clubs and a classroom group. This hybrid virtual community established in order to explore a new form of educational system in order to meet the needs of society and individuals in the 21st century. We engaged with a wide variety of perspectives on educational practices, consistently enacting a view that genuine participation by learners must be instantiated at all stages of education. Within the community, technology is seen not only as a tool to support and extend existing practices but also as having the potential to transform ways of representing the world and of supporting learning. The community decided to explore the potential of virtual worlds, considering their capacity to act as spaces in which visions of future practices and pedagogies can be built and experienced, making it "possible to construct, investigate, and interrogate hypothetical worlds," (Squire, 2006, p. 19) and received funding for three phases of work using the 3D virtual environment Teen Second Life establishing the first 'protected island' in Europe (Gillen et al., 2009).

Considering learning in an ecological perspective entails a broadening out of an investigation into a specific activity beyond its temporal and spatial boundaries (Barron, 2006: 193). The idea of ecology alerts us to a holistic sensitivity, an approach to literacy research that is most often characterised as ethnographic (Barton, 2007; Steinkuehler, 2007); involving a mixture of methods designed to better elucidate the multiple perspectives of participants. Ethnographic study of a virtual world can be approached through bringing in tried and tested methods from social anthropology (Boellstorff, 2008). I propose here that there are also opportunities for developing new methods and concentrate here on discussing one. This then is only a small piece of the jigsaw then that would be required to assemble the holistic approach that an allegiance to ecology demands.

Operating in a virtual world project entailed complex new literacies (Lankshear & Knobel, 2006; Tusting 2008). We communicated using a variety of semiotic resources in multimodal domains: including 'inworld' (to use the prevailing term of activities in the virtual world itself, when projected through avatars), asynchronous discussion fora and collaboratively authored wiki, as further discussed elsewhere (Gillen, 2009). Since participants did not interact outside the project – I never met any students during its duration nor even knew any of their names – learning was only visible through these textually-mediated practices.

The notion of ecology provides a useful metaphor to assist us in thinking about "how the activity – literacy in this case – is part of the environment and at the same time influences and is influenced by the environment" (Barton, 2007, p. 29). Unpicking

the notion of environment in this case is, as ever, a complex matter. Taking a sociohistorical perspective, I consider that the environment to be dynamic, actively constituted by meaning-making processes that the participants engage in, during social interactions, influenced by practices of other domains. So, for example, whether or not people had any previous experience of virtual worlds, participants might draw upon console or computer based gaming experiences, and it was evident that sometimes people did just that. My interest in the project lies in Scheme Park as a learning community; the SPP deriving its name originally in a slogan of 'Scheme – not school not home' was over some years involved with rethinking aims and methods of learning. The project sought explicitly to challenge certain deep-seated practices and assumptions concerned with school-based learning. In some participants' vision Scheme aimed to offer a new model of how traditional schooling could be overthrown. However, at the very same time the project constantly engaged with schools and school-based students. My own outlook was perhaps initially to consider the project as a brave endeavour to bridge the widely recognized 'gap' between school and out-of-school literacy practices. However, I have come to think that a learning ecology perspective may well provide a more useful way of thinking about a challenge that certainly exists than the dichotomy inherent in the notion of 'gap' or 'divide'.

It is a characteristic of human learning that we do make connections between activities in our different domains. Children are very adaptable, and may well, as they have done throughout history, embody very different practices of behaviour and activity in schools than they do outside. But success in such learning requires as a prerequisite the ability to recognise, however unconsciously or consciously, the differences in those cultural environments, made up as they are of distinct material differences and practices characteristic to each (Cole, 1996). One learns to act, as a social human being, the appropriate ways of interacting with each environment. Learning can be regarded then as transformations in the patterns of participation in joint activity (Rogoff, 1997). Although participation in the SPP was diverse according to any possible measurement, a quality that everybody shared was that it was a new experience for everybody, innovative in terms of environment, community and aims, thus demanding of new practices and positioning everybody, in at least some aspects of their practices, as learners.

My aim in this paper is to explore the explicit links participants made between schooling and the project, using methods of discourse analysis, specifically corpus linguistics, to investigate an enormous dataset. My exploration is both substantive and methodological: I seek to contribute to this symposium in two ways: to enhance our understandings of the ways in which students made connections between learning in two very different domains and also to present possibly new methodological tools in the shape of techniques of corpus linguistics. I need first to bring in a necessarily extremely brief introduction to the SPP.

The Schome Park Programme

The project ran from March 2007 to May 2008 in several phases that were differentiated in varying degrees in terms of participation, objectives and activities. At the same time there was considerable continuity in terms of setting, core participants, general ethos and some recurring activities. Appendix A is an attempt to summarise the project drawn from Twining & Footring (2008). A summary overview of the project has been published (Twining, 2009). The aims of the leadership of the project evolved as shown in Appendix A – although it is debatable to what extent any shift in aims was shared amongst all members of the community (Gillen, 2010). Teenagers participating in the SPP were also in almost every case also attending school (although there were a few exceptions who were home-schooled or on long term sickness leave). Some participated as part of their school activity, although as a distinct project; some through an after school club. For every student participating, joining the SPP was voluntary; for the majority it was a home-based activity which they freely chose to join after hearing of the project, typically through information made available via schools or the National Association for Gifted and Talented Youth. This latter organisation was the recruiting channel in the pilot phase, providing a small core of students that remained active and influential throughout as many other students joined from the UK and USA. Patterns of staff involvement were also diverse. The project had an enormous range of activities during the period, some connected with formal curriculum topics, e.g., physics, ethics and philosophy, and archaeology. Other activities stemmed from playful exploration of the environment's affordances and attempts to share new skills with other participants. Many community games, events, and activities were spawned, some initiated by staff and some by students, instantiating fluid leadership (Peachey, Gillen & Ferguson, 2008).

Such a large-scale project has the capacity to generate enormous datasets. Before explaining the approach illustrated here, I briefly outline why corpus linguistics has not yet been taken up to any great degree in discourse analysis in education.

Corpus linguistics

In my opinion, the tools of corpus linguistics have not yet been taken up in education research to the degree they merit owing to two inter-related factors. One of these is concerned with the history of corpus linguistics – that has, understandably enough, shaped the concerns, aims and thus of course the practices and results typical of its practitioners. The majority of work in corpus linguistics has been and continues to be concerned with large-scale investigations of patterns in language in order to find out more about language use at the macro scale, out of immediate consideration of specific features of temporal and spatial context. As

valuable as this work and its applications are, they are not likely to be core to the interests of literacy educationalists except for those concerned with second language learning, who can make use of its consequent pedagogical resources (usage based dictionaries etc.). The second factor as to why the potential contribution of corpus linguistics may be overlooked lies in the way it is sometimes presented in contrast with discourse analysis through its deployment of quantitative methods (e.g. Tognini Bonelli 2010: 19). If a literacy researcher employs qualitative methods they are likely to recognise the worth of discourse analysis, that is the value of taking a sustained detailed look at a small stretch of text and perhaps have immediate doubts as to what role quantitative methods can have to interpretive work.

Nevertheless, the potential of corpus linguistics methods to assist in discourse analysis is being increasingly advocated (Baker, 2006: McCarthy & O'Keefe, 2010). Such a stance is consonant with my own sense of corpus linguistics as existing within the broad umbrella that is discourse analysis (Gillen & Petersen, 2005). So, if the heartland of corpus linguistics is based on definitions such as 'the study of language based on examples of real language use' (McEnery & Wilson, 1996), thus putting the focus on 'the study of language', it is nevertheless becoming clear that there are opportunities for those whose aims are associated with the exploration of specific discourses, and/or specific phenomena that generate large textual datasets, to make use of its methods. In this paper I offer my own introduction to techniques of corpus linguistics through demonstration of my response to the substantive question.

A corpus linguistic investigation of the Schome Park chatlogs

During the project staff members periodically archived some of the chatlogs they had collected during the project. These represented a small part of the records of the interactions inworld, mostly virtually synchronous with some instant messages. With the exception of the resources collected by one staff member (for ethical reasons) these have been collated and organised into a corpus of data. This essentially means that they have been converted into a shared format (.txt), named in a consistent fashion and generally organised in order to be useable by Wordsmith Tools (Scott, 2008). Initial analytic work as describes below ascertains some core facts about the corpus, while introducing some basic corpus linguistics tools.

The Schome Park corpus consists of 682 files contributed by 22 people. A total of 2,443,495 tokens, ie running words in the texts can then be used for word lists and other operations. For example, it is easy to produce a list of the most common words appearing in the corpus – see fig. 1 below.

N	Word	Freq.	%	Texts	%_lemmas	Set
1	#	815,078	25.01	666	97.65	
2	SCHOMER	156,561	4.80	558	81.82	
3	YOU	79,938	2.45	605	88.71	
4	TO	58,735	1.80	597	87.54	
5	THE	51,302	1.57	605	88.71	
6	I	45,937	1.41	584	85.63	
7	A	39,582	1.21	547	80.21	
8	SENSOR	35,520	1.09	38	5.57	
9	PRIM	35,092	1.08	116	17.01	
10	OBJECT	32,020	0.98	198	29.03	
11	IN	29,001	0.89	542	79.47	
12	IS	26,542	0.81	541	79.33	
13	IT	24,553	0.75	542	79.47	
14	AND	22,820	0.70	551	80.79	
15	SPARKER	21,992	0.67	270	39.59	
16	OF	18,816	0.58	487	71.41	
17	TEST	18,370	0.56	89	13.05	
18	THAT	17,657	0.54	485	71.11	
19	BE	16,388	0.50	503	73.75	
20	AT	15,863	0.49	419	61.44	

frequency alphabetical statistics filenames notes

34,169 Type-in #

Figure 1: The top 20 words in the Schome Park chatlog corpus, presented in order of frequency.

The frequency column indicates the total number of times the word occurs in the texts; the next reflects the proportion of that word in the corpus; 'texts' indicates the number of texts within the corpus the word appears in and the proportion this represents. I will return to the blank column 'lemmas' below. A few remarks on the findings as to these common words can be made in order to point out some distinctive features of the discourses of this environment:

- # means a number or a word that includes a number
- SCHOMER and SPARKER are both words that denote project participants, appear automatically on the log as the avatars' surnames and thus have not been generated each time by participants (in the way that words they personally type are generated).
- SENSOR, PRIM, OBJECT are all words that have specific meanings in the environment. For example a prim is a basic building block. Objects, which have been constructed within the environment are often

programmed ('scripted' in Second Life terms) to automatically emit a message when 'touched' by an avatar. So, for example, one might easily encounter in a log 'Object: you aren't the owner' – this is the result of scripting prior to the interaction.

With so many scripted i.e. preprogrammed words, it is clear that the corpus is unlike most corpora that are made up of utterances and texts that are crafted individually (I mean utterances and texts that are produced at a specific time and place, recognising that many texts such as newspaper articles are the work of more than one author). The Schome chatlog corpus then features a mixture of language that is not generated spontaneously and a great deal of repetition.

An often useful calculation that is made is that of the type/token ratio (TTR) answering the questions how many words are in a corpus, how diverse is its vocabulary? With 34,169 types (distinct words) in the Schome Park corpus the TTR is 1.40. In order to try to compare this with other corpora it is sensible to take a measure known as the standardised TTR (otherwise the larger a corpus gets the more highly frequent grammatical words like *the* will dominate); for this corpus it is 23.30. This figure shows the influence of the high proportion of scripted words; in comparison a standardised TTR of the informal spoken conversations from the British National Corpus, completely without bots (objects scripted to interact with avatars) and indeed avatars was 32.96 (Baker, 2006: 52).

Corpus linguistic studies often re-examine such a frequency list by excluding grammatical words -all the generally short conjunctions, pronouns, etc. that are the 'mortar' rather than the 'bricks' of language; I have constructed a top 10 of lexical words as Table 1:

	Word	Frequency	%	texts	%
1	sensor	35,520	1.09	38	5.57
2	prim	35,092	1.08	116	17.01
3	object	32,020	0.98	198	29.03
4	test	18,370	0.56	89	13.05
5	do	11,009	0.34	467	68.48
6	IM	10,439	0.32	325	47.65
7	can	10,206	0.31	475	69.65
8	P	9,907	0.30	297	43.55
9	sense	9,067	0.28	125	18.33
10	whispers	8,478	0.26	186	27.27

Table 1: Top lexical words, excepting proper nouns, in the Schome corpus

I can bring in my ethnographic perspective as a community member to further emphasise the large part scripted language is contributing to the flow of interaction as recorded here: for example many bots were programmed to *whisper* their turns, neatly downplaying the affective quality of their interruptions. On the other hand as far as I know *P* formed part of emoticons generated by human participants in the flow of communications. I have screened proper nouns, other than the shared surnames, for this paper in part for ethical reasons. Although clearance to use data from contributing participants has been obtained, I have preferred here to anonymise details that would convey both general levels of participation in the project and specific details of turns in ways that would be meaningful to project participants.

The general characterisation made so far served only to introduce some very general features of the environment, rather than to directly further the main aim, to examine connections made between schooling and learning in the Shome Park environment. One way this might be investigated through some corpora is the investigations of collocations: for example with which words does *school*, for example, most often occur? Then those occurrences could be investigated within texts in order to ascertain more of the immediate textual context (Evison, 2010). However, experimental investigations along these lines drew my attention to some difficulties with this.

Since members of staff were often inworld together, their deposited logs often overlapped to some extent. This was never a precise overlap for a number of reasons including the duration of stay, proximity to other avatars etc. Nevertheless it did strike me as somewhat problematic, as potentially it could make attach too much weight to some discourses. (There is an alternative reading however, that if an interaction did feature high participation, for example at a well-attended meeting, then its relatively high impact would be reflected through multiple deposits). Further investigating overlaps, I found that some people had submitted chatlogs that themselves contained overlaps with a previously submitted file.

It is a very common practice in corpus linguistics to decide to deploy a sampling strategy (Biber, Conrad & Reppen, 1998; Baker, 2006). The researcher is trying to seek a balance between creating a dataset that is manageable, within the constraints of the particular exercise, while still utilising sufficiently large amounts of data. In this case after some experiment I decided to sample the largest single files collected by four individuals – two male and two female. Through an iterative process, I eventually decided to search the files for instances of four lemmas. Lemmas are related word forms: for example if a search of the files was only centred on *school* and not *schools* clearly a great deal of relevant information might be lost. Many lemmas can be captured through a Boolean strategy of *school** but not all. For example, I searched

for *teach** but also *taught*. *Class* often featured in my sample but *class** cannot be used as a search strategy since it picks up *classed classify* etc. So, I used *class, classes classroom* and *classrooms*, but then examined all instances of *class* to ensure it was not used in another sense, i.e. to exclude homonyms. The extremely rapid typing common inworld brought about many misspellings and of course as far as possible these also needed picking up, e.g. *schoool*.

Analysing the largest file first, I discovered some basic quantitative measurements, compiled the four lemmas, conducted searches for them, recording ranks and frequencies. I then searched all the concordances (as explained below) to further investigate the texts in all four files where these lemmas occurred. Working qualitatively I grouped these according to themes I identified. For reasons of space I present here the findings relating to two of the files.

GB_009

Staff member GB file_009 has 917.782 tokens; 406,603 used for word list. The standardised TTR is 5.29 which is almost extraordinarily incredibly low. A brief look at this log revealed an extremely technical orientation with # at 55.70% and *sensor* the most common word.

lemma	rank	freq
SCHOOL	416	36
TEACH	2315	33
LEARN	651	24
CLASS	1822	4

Table 2: GB_009: results for investigated terms

Table 2 then shows the results of the searches for targeted lemmas in file GB_009. By convention in corpus linguistics lemmas are shown in small caps. The next part of the exercise is to locate these terms in context; this can be done through the software's concordance function. Figure 2 shows an extract from a sample concordance view, in which names that would be identifiable to project participants have been erased.

5	like you have teams for newspapers at school [2007/11/19 13:20]
6	Schomer: i bought a smoothie at school and I went to open it, I sliced right
7	11:54] ... at school we haven't even been back for two
8	SParker: had a bad day at school and went lookign for trouble
9	SParker: has had a bad day at school so looked to stir trouble here
10	a long long while since you were at school [2007/11/19 13:43]
11	it's hard when you're crossing school with second life, where the ethos
12	hard to directly talk with during school hours [2007/11/19 13:23]
13	SParker: I teach in the elementary school , [2007/11/19 13:22]
14	12:55] : the entire school year! [2007/11/19 12:55]

Figure 2, extract from concordance file for GB_009 for school*

The data in Figure 2 has been alphabetically sorted first one place to the left and then second place to the left. The next step is to click through to locate each of these samples in the surroundings of its immediate text, in order to enable interpretation. I was then able to inductively derive themes with which to categorise all findings in the file. With examples of varying length I present those for this file below.

SCHOOL

Schools are sometimes described in terms of their associations with regimes and rules, places where you have to do something, e.g. "LB Schomer: but, in school, we had to stand when...." or "I admit I found the whole school regime uneducational". In the course of a complex discussion of the rules the Schome community created to govern themselves is found: "...rules might revert to 'normal school conduct')?"

School time and temporal organisation sometimes feature, for example in reference to a group of participants: "the UK who will be using it in school time." One student suddenly interjected in the course of a communication about another topic, "school tomorrow." EF Schome manipulated the usual convention of producing a turn as if uttered by an avatar into a third person statement about himself: "EF Schomer has to get up for school at around 7. The evilness!"

A variety of RL (real life) school experiences were alluded to, such as "I did that at secondary school." Although all student participants were teenagers, mentions were made of other kinds of school including elementary and summer school.

In this dataset praying at school was occasionally referred to: "we dont pray at school.... unkes you go to C..." and "we have prayers at school ever day."

I devoted more detailed attention to explicit connections and contrasts between school and the SPP as first evidenced in three turns:

"FE had a bad day at school and went lookign for trouble." This was staff member TJ explaining that a student had come inworld with a bad temper and decided to take over some land, build, swear and insult others, especially one other. This led to a lengthy discussion in which it was clearly understood by all participating that although you are represented by a distinct avatar you cannot necessarily act as a *tabula rasa*, immune to events outside.

"it's hard when you're crossing school with second life, where the ethos of both is completely different." Here student US is explaining a perceived difficulty caused by one group of students joining the project and coming in with their teacher, whom they actually see every day – and that this seems odd to the main body of students for whom the project has nothing to do with their RL school staff. A staff member SB offers various suggestions such as providing areas where staff cannot go (an idea immediately resisted by student LB). SB questions "whether we can change the way in which teachers view their role ... or is that only possible where they are working with students who they don't have to work with face to face?" US responds: "i think it would be much harder where the teachers see the students in RL as well," claiming there are unwritten rules about the way students act around teachers in school. The discussion moves to whether new students could be supervised by other staff – LB suggests other experienced students could do this.

"...like you have teams for newspapers at school..." Here SB, trying to improve SPP communications suggests a newsletter and proposes the role of a newsletter coordinator. A student immediately makes a connection to a similar kind of activity organisation experienced at school.

TEACH

As already indicated above, there are references to RL teachers in the corpus, sometimes it seems just as part of a small narrative whereby students bring in RL events, e.g. "my cornet teacher shes scary at times" and elsewhere "but I have a more relaxed time with my teachers now that i'm at college".

Teaching activities within the project are announced. "scripting and have activities teaching real life physics and chemistry" meant, (perhaps slightly confusingly for us at this distance,) that within SPP actual physics and chemistry were going to be taught.

Pertinent for this study are reflections on the SPP ethos, whereby everybody was seen as a learner and teacher at various times. Student US: "i like the way that everyone has different skills, and can teach everyone else things, even the adults, who would be seen as teachers in RL and therefore the person whom people go to for help and things."

It would of course at the same time be naïve to expect that with the staff role did not come responsibilities and therefore a certain recognition of power relations. During the discussion cited above about the dilemma perceived as being caused by the group of students who were accompanied inworld by a RL teacher one student appeared to turn a little impatient with TB's attempt to bring about a consensual solution. EB: "well u are our teacher ys".

LEARN

I categorised the 24 instances in this file into three themes. First were references to Second Life skills, such as: "now all I need to learn is how to put it in a loop and not..." "i think that one of the easiest scripts you can learn is the prim shape and colour change script.."

There were positive statements about learning in the SPP: "we probably learn as much from you as you learn from us." In a contrast made with patterns of learning in SL a student suggested there is "a better feeling that everyone is a learner together" [in the SPP].

'Learn' was also used in a very general way referring to learning from life experience. GB: "you will soon learn that EVERY SINGLE government comp scheme is crap.

CLASS

References were made to environments where instructional events were held in the SPP, such as "what brings you to our classroom-turned-cemetery?" and connections made with learning elsewhere, "reminds me of a class I did on MG on animations and incorporating gestures..."

SB_154

This staff member's few technical responsibilities and broad diversity of interests are immediately suggested by some basic facts of the file: it has 98,015 tokens of which 96,068 were used for the word list. His standardised TTR was 33.34 which since it includes scripted words can be judged to be indicative of a very broad vocabulary. A comparison of the ranking in this file of the target words with those of Table 2 reflect this staff member's interest in discussing these topics. The constitution of lemmas demonstrates some linguistic creativity; for example TEACH includes *teacher teachers teachery teacheryness teachinf teaching teachrs*

lemma	rank	freq
SCHOOL	194	88
TEACH	1079	85
LEARN	499	55
CLASS	526	32

Table 3 SB_154: results for investigated terms

As explained above, here I add to findings in indicative new ways rather than citing examples that support the themes discussed above. In respect of SCHOOL I mention one more aspect of use discovered through study of concordances and finally proceed to a brief analysis of a stretch of discourse related to LEARNING.

SCHOOL

There were extensive discussions of schooling, its organisation and setting into tiers, learning styles etc. Topics new here included bullying, eg. NJ "yes true but it's mainly you're bullied (at my school) for being smart or looking like a goth. i'm smart, a black sheep and i like black clothing."

Discourse analysis: LEARNING in file SB_154

Following is a stretch of discourse that contains both 'learning on the go' as it proceeded in this collaborative community and reflections on learning. I have added turns for the ease of reference. The passage begins where BD Schomer has just been asked by a member of staff about what characterises learning in SPP.

1. BD Schomer: laissez-faire really, theres no formality most of the time
2. EF Schomer: Indeedie.
3. UP Schomer: I were the cloack of a lord now peasent viba:P
4. FN Schomer: cool cape UP
5. EF Schomer remembers when he learnt how to not splat on the floor when he stopped flying :p
6. EF Schomer: That was quite random :p
7. WJ Schomer: that took me ages i kept having to click pgdown
8. [Turn omitted for ethical reasons]
9. FN Schomer: how do u stop splatting when u stop flying
10. FN Schomer: ?
11. EF Schomer: One thing about the way of learning stuff here is that, unlike in school, you can mess around a lot more
12. WJ Schomer: and i couldn't go up or down... i cant remeber who told me about the pgscroll thing
13. BD Schomer: a lot more experimentation
14. EF Schomer: So if you want to test something, you can; in school, you have to stick to what you're told to do...

15. WJ Schomer: FN just press pg dn
16. WJ Schomer: i agree dec... its not rigid here
17. EV Schomer: I'm afraid i have to leave now
18. EV Schomer: a music lesson awaits!
19. EF Schomer: bye then
20. UP Schomer has to go to finish th final copy of history
21. EV Schomer: goodbye!
22. WJ Schomer: okee byee
23. BD Schomer: bye

A distinctive feature of the discourse here is its fast pace and multiple weaving of topic threads. The thread shown initially in turn 1 is constituted by reflections on learning in the SPP; BD remarks upon its informality, this is picked up by EF in turn 11 who mentions messing around. That such a term is perhaps used approvingly is suggested by BD's more specific response, "a lot more experimentation" in turn 13. This is taken up immediately by EF who refers to SPP as somewhere where "if you want to test something, you can" and contrasts this with school. I find it delightful though that this dialogue is interwoven with indications of a specific learning event. It becomes apparent through reading the transcript that EF's recollection of learning a skill: "EF Schomer remembers when he learnt how to not splat on the floor when he stopped flying :p" is, it seems, a tactful suggestion to FN who has presumably just committed this 'newbie' action. Given the time it takes to produce a turn, FN probably immediately types out his request for the relevant knowledge at turn 9 with a direct question. WJ, again with some tact, has actually already imparted the answer, but in case FN has not picked this up later responds directly in turn 15. Her use of his name also attends to his needs; it took time for a newcomer to adjust to the fastflowing multiply threaded dialogues and in such a situation to be addressed directly can help. Finally, RL learning events and tasks are drawn upon to bring about a polite end to the discussion, indicative of participants' sense that they belong to multiple learning worlds and are committed to the constructive of collaborative social ties in this one, even with associates they will never meet face to face. .

Conclusions

This work has shown in practice that when dealing with an enormous dataset it can be productive to utilise tools belonging to corpus linguistics, considering this as part of a larger umbrella of discourse analysis activity (Gillen & Petersen, 2005). My approach is intended to contribute to the breadth of approaches adapted from linguistics to studies of multimedia learning such as that for example offered by Steinkuehler (2006) in her application of functional linguistics to a study of MMOG online gaming. This seems ultimately more fruitful than either dichotomising corpus linguistics and discourse analysis along an outworn quantitative/qualitative dichotomy or trying to construct some sort of artificial barrier based on the size of

texts under consideration (an approach effectively dismissed by Baker, 2006). At the same time, the use of any tools from linguistics always demand thinking through as to their appropriacy to the 'logic of inquiry' (Green, Dixon & Zaharlick 2003).

Drawing on the work of Bransford and Schwartz, Barron, (2006:221) argues that educational experiences should be evaluated according to their potential for providing students with the opportunities to learn, in a highly social setting with plenty of feedback to generate what she terms 'self-sustaining learning ecologies.' This sociocultural orientation to learning, through activity undertaken with others in spaces imbued with the traces of earlier human cultural activities, is enunciated by Stetsenko (2009, 126):

According to this vision, human development is rooted in, derivative of, instrumental in, and constituted by the material collaborative social practices of people (i.e., human goal-directed, purposeful, collaborative activities) aimed at transforming their world.

For me the notion of a world transformed was experienced as more than a metaphor as I, like many others, learned so much in our immersive new environment. Changes in our social identities permeate more broadly than through the online community in which they are expressed (Merchant, 2006). This paper has not been an exercise in objective evaluation, but rather investigates what were to me and some other participants some of the most interesting aspects of SPP. In a challenging environment adult staff members were often behind students in their grasp of skills (this may be partially accounted for by unequal amount of time spent practising) and, as in many informal learning activities, conventions of traditional teacher/learner relationships were sometimes at least reshaped for the better. Fostering a learning disposition is perhaps the key task for education, whether formal or informal and whatever the age, stage or status of the learner.

Contact: j.gillen@lancaster.ac.uk

References

- Baker, P. (2006) *Using corpora in discourse analysis*. London: Continuum.
- Barron, B. (2006) Interest and self-sustained learning as catalysts of development: a learning ecology perspective. *Human Development* 49 193-224.
- Barton, D. (2007) *Literacy: An introduction to the ecology of written language*. 2nd edn. Oxford: Blackwell.
- Biber, D., Conrad, S. & Reppen, R. (1998) *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

- Boellstorff, T. (2008) *Coming of age in Second Life*. Princeton: Princeton University Press.
- Cole, M. (1996) *Cultural psychology: A once and future discipline*. Cambridge, MA: Belknap Press of Harvard University Press.
- Evison, J. (2010) What are the basics of analysing a corpus? In A. O'Keefe & M. McCarthy (eds) *The Routledge Handbook of Corpus Linguistics*. London: Routledge.
- Gillen, J. (2010) New literacies in Shome Park. In A. Peachey, J. Gillen, D. Livingstone & S. Robbins, (Eds.) *Researching Learning in Virtual Worlds*. Springer.
- Gillen, J. (2009) Literacy practices in Shome Park: a virtual literacy ethnography. *Journal of Research in Reading* (special issue on literacy and technology) 32 (1) 57-74.
- Gillen, J. & Petersen, A. (2005) 'Discourse Analysis' in B. Somekh & C. Lewin (eds) *Research Methods in the Social Sciences*. Thousand Oaks, CA/London: Sage Publications.
- Gillen, J., Twining, P., Ferguson, R., Butters, O., Clough, G., Gaved, M., Peachey, A., Seamans, D. & Sheehy, K. (2009) [A learning community for teens on a virtual island - The Shome Park Teen Second Life Pilot Project](#). eLearning Papers no. 15: The New Learning Generation.
- Green, J.L, Dixon, C.N., & Zaharlick, A. (2003). Ethnography as a logic of inquiry. In J. Flood, D. Lapp, J Squire, J. & J. Jensen (Eds.), *Handbook of research on teaching the English language arts* (2nd ed., pp. 201-224). Mahwah, NJ: Erlbaum.
- Guo, L., Amasha, S & Tan, L. (2011) Multimodal literacy in extended learning activities. In C. Ho, K. Anderson & A.. Leong (eds) *Transforming literacies and language: Multimodality and literacy in the new media age*. London: Continuum.
- Lankshear, C. & Knobel, M. (2006) *New Literacies: everyday practices and classroom learning 2nd edn*. Maidenhead: Open University Press.
- McCarthy, M. & O'Keefe, A. (eds) (2010) *The Routledge Handbook of Corpus Linguistics*. London: Routledge.
- McEnery, T., and Wilson, A. (1996) *Corpus Linguistics*, Edinburgh. Edinburgh University Press.
- Merchant, G. (2006) Identity, social networks and online communication. *E-Learning* 3 (2) 235-244.
- Peachey, A., Gillen, J. & Ferguson, R. (2008) [Fluid leadership in a multi-user virtual environment educational project with teenagers: Shome Park](#). Paper presented at Ecologies of Diversities: the developmental and historical interarticulation of human mediational forms: meeting of the International Society for Cultural and Activity Research, San Diego, USA. September 8-13.
- Rogoff, B. (1997) 'Evaluating Development in the process of participation: theory, methods and practice building on each other' in E. Amsel and K.A. Renninger, (eds). *Change and development: Issues of theory, method and application*. Mahwah, NJ: Erlbaum.

- Scott, M., 2008, WordSmith Tools version 5, Liverpool: Lexical Analysis Software. <http://www.lexically.net/wordsmith/index.html> accessed 15 March 2011.
- Squire, K. (2006) From Content to context: videogames as designed experience. *Educational Researcher* 35(8) 19-29.
- Steinkuehler, C. (2007) Massively multiplayer online gaming as a constellation of literacy practices. *E-Learning* 4 (3) 297-318.
- Steinkuehler, C. (2006) Massively multiplayer online video gaming as participation in a discourse. *Mind, Culture and Activity* 13 (1) 38-52.
- Stetsenko, A. (2009) Vygotsky and the conceptual revolution in developmental sciences: towards a unified (non-additive) account of human development. In M. Flear, M. Hedegaard & J. Tudge (Eds.), *World yearbook of education 2009. Childhood studies and the impact of globalization: Policies and practices at global and local levels* (pp. 125-141). New York: Routledge.
- Tognini Bonelli, E. (2010) Theoretical overview of the evolution of corpus linguistics. In A. O'Keefe & M. McCarthy (eds) *The Routledge Handbook of Corpus Linguistics*. pp. 14- 27. London: Routledge.
- Tusting, K. (2008) Ecologies of new literacies: implications for education. In A. Creese, P. Martin & N. Hornberger (eds) *Encyclopedia of Language and Education 2nd edn. volume 9. Ecology of Language*. Dordrecht. Springer. pp. 317-329.
- Twining, P. (2009) Exploring the educational potential of virtual worlds – some reflections from the SPP. *British Journal of Educational Technology* 40 (3) 496-514.
- Twining, P. & Footring, S. (2008) The Scheme Park Programme – exploring educational alternatives. *Researching Learning in Virtual Environments 2008: The Open University, Milton Keynes, November 2008*. http://www.open.ac.uk/relive08/documents/ReLIVE08_conference_proceedings_Lo.pdf (4MB file accessed 15 March 2011).

Acknowledgements

This paper is embedded in the work of the Scheme Community – regrettably too many people to name. The project was led by Peter Twining of the Open University; people who have particularly helped my research in Scheme Park also include Rebecca Ferguson, Anna Peachey and student participants. I am also grateful to Costas Gabrielatos and Andrew Hardie of Lancaster University for assistance in organising a huge and chaotic dataset into a corpus and to Paul Baker for inspiration through his work. None of the above are responsible for any deficiencies in this foray into a disciplinary territory where I'm very much a learner.

ⁱ Teen Second Life and Second Life are trademarks of Linden Research, Inc.

Appendix A: Overview of Scheme Park project.

Aspect	Phase 1 (March to April 07)	Phase 2 (June to December 07)	Phase 3a (January to March 08)	Phase 3b (April to May 08)
Our focus (Aims)	To explore the educational potential of virtual worlds (with a particular focus on developing Second Life skills and ‘Knowledge Age Skills’) To build a community of learners	To enhance ‘Knowledge Age Skills’ To increase student control and responsibility for the environment, the curriculum and support To widen the community (not just gifted and talented)	To enhance ‘Knowledge Age Skills’ To balance control and responsibility for the environment, the curriculum and support. To widen the community and increase its size To explore the co-existence of the Scheme ethos with school culture	
Environment	Island divided into six areas: Physics Ethics & philosophy Archaeology Scho-op (generic support) Shared meeting areas Sandbox Island, wiki and forum available 24/7/365	Island as naturalistic and attractive environment with some core generic areas – student control of planning/building	Two islands: One student controlled + one staff controlled. Immersive game theme for new island.	Two islands. Project teams allocated plots of land with full controls (e.g. terraforming)
Actors	149 students aged 13 to 17, from the National Association of Gifted and Talented Youth (NAGTY) Staff from four universities Staff from the National Physical Laboratory PhD students Consultants	Ongoing students from Phase 1 New 13 to 17 year old students from range of sources (inc USA) Staff from two universities PhD students Consultants Teachers Parents	Ongoing students from Phase 2 New 13 to 17 year old students from range of sources, including: South East Grid for Learning (broadband consortium) and ‘School groups’ from UK and USA	
Curriculum	Three strands of formal activity (Physics, Ethics and Philosophy, Archaeology) Discrete ‘taught sessions’ (e.g. research methods) Student led activity	Student led activity (inc continuation of formal strands from Phase 1) Machinima creation Discrete ‘staff led’ sessions (e.g. Sudoku)	Student led activity (inc continuation of Phase 2 strands and new strands such as Time Travellers) New strands led by staff (e.g. Math’s)	Major focus on projects (led by students and/or staff)
Support	Staff scheduled sessions for each formal curriculum area Peer – peer support Information in wiki Discussion in forum Emergency help button to summon staff	Staff available to provide support in Scheme Park	Greater staff support for strands of activity (e.g. Math’s). Greater support for student led activity	Staff support focused on projects